

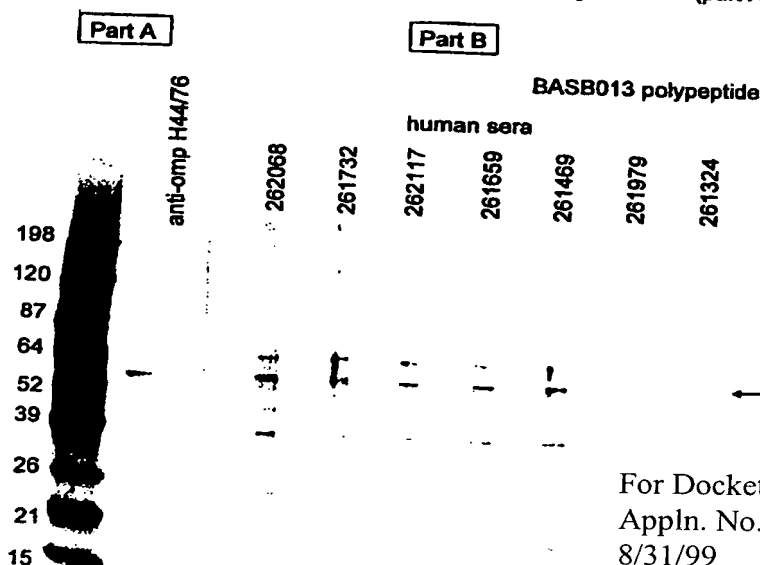


INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 : C12N 15/31, 15/11, 15/70, 1/21, C07K 16/12, G01N 33/53, 33/50, A61K 31/70, 39/095, 39/40, C07K 14/22		A1	(11) International Publication Number: WO 99/55872
			(43) International Publication Date: 4 November 1999 (04.11.99)
(21) International Application Number: PCT/EP99/02765		(81) Designated States: AE, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZA, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).	
(22) International Filing Date: 20 April 1999 (20.04.99)			
(30) Priority Data: 9808734.9 23 April 1998 (23.04.98) GB			
(71) Applicant (for all designated States except US): SMITHKLINE BEECHAM BIOLOGICALS S.A. [BE/BE]; 89, rue de l'Institut, B-1330 Rixensart (BE).			
(72) Inventor; and			
(75) Inventor/Applicant (for US only): RUELLE, Jean-Louis [BE/BE]; SmithKline Beecham Biologicals S.A., 89, rue de l'Institut, B-1330 Rixensart (BE).			
(74) Agent: TYRRELL, Arthur, William, Russell; SmithKline Beecham, Two New Horizons Court, Brentford, Middlesex TW8 9EP (GB).		Published With international search report. Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.	

(54) Title: BASB013 DNA AND PROTEINS FROM NEISSERIA MENINGITIDIS

Anti-BASB013 antibodies in human convalescent sera (part B) and mice immunized with Outer Membrane Proteins of H44/76 Neisseria meningitidis cells (part A).

For Docket No.: 7969-082-999
Appln. No. 09/388,090 filed
8/31/99

Reference: AH

(57) Abstract

The invention provides BASB013 polypeptides and polynucleotides encoding BASB013 polypeptides and methods for producing such polypeptides by recombinant techniques. Also provided are diagnostic, prophylactic and therapeutic uses.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece			TR	Turkey
BG	Bulgaria	HU	Hungary	ML	Mali	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MN	Mongolia	UA	Ukraine
BR	Brazil	IL	Israel	MR	Mauritania	UG	Uganda
BY	Belarus	IS	Iceland	MW	Malawi	US	United States of America
CA	Canada	IT	Italy	MX	Mexico	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NE	Niger	VN	Viet Nam
CG	Congo	KE	Kenya	NL	Netherlands	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NO	Norway	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	NZ	New Zealand		
CM	Cameroon			PL	Poland		
CN	China	KR	Republic of Korea	PT	Portugal		
CU	Cuba	KZ	Kazakstan	RO	Romania		
CZ	Czech Republic	LC	Saint Lucia	RU	Russian Federation		
DE	Germany	LI	Liechtenstein	SD	Sudan		
DK	Denmark	LK	Sri Lanka	SE	Sweden		
EE	Estonia	LR	Liberia	SG	Singapore		

BASB013 DNA AND PROTEINS FROM NEISSERIA MENINGITIDIS

FIELD OF THE INVENTION

This invention relates to polynucleotides, (herein referred to as "BASB013 polynucleotide(s)"), polypeptides encoded by them (referred to herein as "BASB013" or "BASB013 polypeptide(s)"), recombinant materials and methods for their production. In another aspect, the invention relates to methods for using such polypeptides and polynucleotides, including vaccines against bacterial infections. In a further aspect, the invention relates to diagnostic assays for detecting infection of certain pathogens.

BACKGROUND OF THE INVENTION

Neisseria meningitidis (meningococcus) is a Gram negative bacterium frequently isolated from the human upper respiratory tract. It occasionally causes invasive bacterial diseases such as bacteremia and meningitis. The incidence of meningococcal disease shows geographical seasonal and annual differences (Schwartz, B., Moore, P.S., Broome, C.V.; Clin. Microbiol. Rev. 2 (Supplement), S18-S24, 1989). Most disease in temperate countries is due to strains of serogroup B and varies in incidence from 1-10/100,000/year total population sometimes reaching higher values (Kaczmarek, E.B. (1997), Commun. Dis. Rep. Rev. 7: R55-9, 1995; Scholten, R.J.P.M., Bijlmer, H.A., Poolman, J.T. et al. Clin. Infect. Dis. 16: 237-246, 1993; Cruz, C., Pavez, G., Aguilar, E., et al. Epidemiol. Infect. 105: 119-126, 1990).

Epidemics dominated by serogroup A meningococci, mostly in central Africa, are encountered, sometimes reaching levels up to 1000/100.000/year (Schwartz, B., Moore, P.S., Broome, C.V. Clin. Microbiol. Rev. 2 (Supplement), S18-S24, 1989). Nearly all cases as a whole of meningococcal disease are caused by serogroup A, B, C, W-135 and Y meningococci and a tetravalent A, C, W-135, Y polysaccharide vaccine is available (Armand, J., Arminjon, F., Mynard, M.C., Lafaix, C., J. Biol. Stand. 10: 335-339, 1982).

The polysaccharide vaccines are currently being improved by way of chemical conjugating them to carrier proteins (Lieberman, J.M., Chiu, S.S., Wong, V.K., et al. JAMA 275 : 1499-1503, 1996).

- 5 A serogroup B vaccine is not available, since the B capsular polysaccharide was found to be nonimmunogenic, most likely because it shares structural similarity to host components (Wyle, F.A., Artenstein, M.S., Brandt, M.L. et al. J. Infect. Dis. 126: 514-522, 1972; Finne, J.M., Leinonen, M., Mäkelä, P.M. Lancet ii.: 355-357, 1983).
- 10 For many years efforts have been initiated and carried out to develop meningococcal outer membrane based vaccines (de Moraes, J.C., Perkins, B., Camargo, M.C. et al. Lancet 340: 1074-1078, 1992; Bjune, G., Hoiby, E.A. Gronnesby, J.K. et al. 338: 1093-1096, 1991). Such vaccines have demonstrated efficacies from 57% - 85% in older children (>4 years) and adolescents.
- 15 Many bacterial outer membrane components are present in these vaccines, such as PorA, PorB, Rmp, Opc, Opa, FrpB and the contribution of these components to the observed protection still needs further definition. Other bacterial outer membrane components have been defined by using animal or human antibodies to be potentially relevant to the induction
- 20 of protective immunity, such as TbpB and NspA (Martin, D., Cadieux, N., Hamel, J., Brodeux, B.R., J. Exp. Med. 185: 1173-1183, 1997; Lissolo, L., Maître-Wilmotte, C., Dumas, p. et al., Inf. Immun. 63: 884-890, 1995). The mechanisms of protective immunity will involve antibody mediated bactericidal activity and opsonophagocytosis.
- 25 A bacteremia animal model has been used to combine all antibody mediated mechanisms (Saukkonen, K., Leinonen, M., Abdillahi, H. Poolman, J. T. Vaccine 7: 325-328, 1989). It is generally accepted that the late complement component mediated bactericidal mechanism is crucial for immunity against meningococcal disease (Ross, S.C., Rosenthal P.J., Berberic, H.M., Densen, P. J. Infect. Dis. 155: 1266-1275, 1987).

The frequency of *Neisseria meningitidis* infections has risen dramatically in the past few decades. This has been attributed to the emergence of multiply antibiotic resistant strains and an increasing population of people with weakened immune systems. It is no longer
5 uncommon to isolate *Neisseria meningitidis* strains that are resistant to some or all of the standard antibiotics. This phenomenon has created an unmet medical need and demand for new anti-microbial agents, vaccines, drug screening methods, and diagnostic tests for this organism.

10

SUMMARY OF THE INVENTION

The present invention relates to BASB013, in particular BASB013 polypeptides and BASB013 polynucleotides, recombinant materials and methods for their production. In another aspect, the invention relates to methods for using such polypeptides and
15 polynucleotides, including prevention and treatment of microbial diseases, amongst others. In a further aspect, the invention relates to diagnostic assays for detecting diseases associated with microbial infections and conditions associated with such infections, such as assays for detecting expression or activity of BASB013 polynucleotides or polypeptides.

20

Various changes and modifications within the spirit and scope of the disclosed invention will become readily apparent to those skilled in the art from reading the following descriptions and from reading the other parts of the present disclosure.

25

DESCRIPTION OF THE INVENTION

The invention relates to BASB013 polypeptides and polynucleotides as described in greater detail below. In particular, the invention relates to polypeptides and polynucleotides of

BASB013 of *Neisseria meningitidis*, which is related by amino acid sequence homology to the MucD protein of *Pseudomonas aeruginosa*. The invention relates especially to BASB013 having the nucleotide and amino acid sequences set out in SEQ ID NO:1,3, 5 and SEQ ID NO:2,4, 6 respectively. It is understood that sequences recited in the Sequence Listing below as "DNA" represent an exemplification of one embodiment of the invention, since those of ordinary skill will recognize that such sequences can be usefully employed in polynucleotides in general, including ribopolynucleotides.

Polypeptides

- 10 In one aspect of the invention there are provided polypeptides of *Neisseria meningitidis* referred to herein as "BASB013" and "BASB013 polypeptides" as well as biologically, diagnostically, prophylactically, clinically or therapeutically useful variants thereof, and compositions comprising the same.
- 15 The present invention further provides for:
- (a) an isolated polypeptide which comprises an amino acid sequence which has at least 85% identity, more preferably at least 90% identity, yet more preferably at least 95% identity, most preferably at least 97-99% or exact identity, to that of SEQ ID NO:2,4, 6 ;
 - (b) a polypeptide encoded by an isolated polynucleotide comprising a polynucleotide sequence which has at least 85% identity, more preferably at least 90% identity, yet more preferably at least 95% identity, even more preferably at least 97-99% or exact identity to SEQ ID NO:1, 3, 5 over the entire length of SEQ ID NO:1, 3, 5 respectively; or
 - (c) a polypeptide encoded by an isolated polynucleotide comprising a polynucleotide sequence encoding a polypeptide which has at least 85% identity, more preferably at least 90% identity, yet more preferably at least 95% identity, even more preferably at least 97-99% or exact identity, to the amino acid sequence of SEQ ID NO:2,4, 6 ;

The BASB013 polypeptides provided in SEQ ID NO:2,4, 6 are the BASB013 polypeptides from *Neisseria meningitidis* strains ATCC13090 and H44/76.

The invention also provides an immunogenic fragment of a BASB013 polypeptide, that is, a contiguous portion of the BASB013 polypeptide which has the same or substantially the same immunogenic activity as the polypeptide comprising the amino acid sequence of SEQ ID NO:2,4, 6. That is to say, the fragment (if necessary when coupled to a carrier) is capable of raising an immune response which recognises the BASB013 polypeptide.

Such an immunogenic fragment may include, for example, the BASB013 polypeptide lacking an N-terminal leader sequence, and/or a transmembrane domain and/or a C-terminal anchor domain. In a preferred aspect the immunogenic fragment of BASB013 according to the invention comprises substantially all of the extracellular domain of a polypeptide which has at least 85% identity, more preferably at least 90% identity, yet more preferably at least 95% identity, most preferably at least 97-99% identity, to that of SEQ ID NO:2,4, 6 over the entire length of SEQ ID NO:4.

A fragment is a polypeptide having an amino acid sequence that is entirely the same as part but not all of any amino acid sequence of any polypeptide of the invention. As with BASB013 polypeptides, fragments may be "free-standing," or comprised within a larger polypeptide of which they form a part or region, most preferably as a single continuous region in a single larger polypeptide.

Preferred fragments include, for example, truncation polypeptides having a portion of an amino acid sequence of SEQ ID NO:2,4, 6 or of variants thereof, such as a continuous series of residues that includes an amino- and/or carboxyl-terminal amino acid sequence. Degradation forms of the polypeptides of the invention produced by or in a host cell, are also preferred. Further preferred are fragments characterized by structural or functional attributes such as fragments that comprise alpha-helix and alpha-helix forming regions, beta-sheet and beta-sheet-forming regions, turn and turn-forming regions, coil and coil-forming regions, hydrophilic regions, hydrophobic regions, alpha amphipathic regions, beta

amphipathic regions, flexible regions, surface-forming regions, substrate binding region, and high antigenic index regions.

- Further preferred fragments include an isolated polypeptide comprising an amino acid sequence having at least 15, 20, 30, 40, 50 or 100 contiguous amino acids from the amino acid sequence of SEQ ID NO:2,4, 6 or an isolated polypeptide comprising an amino acid sequence having at least 15, 20, 30, 40, 50 or 100 contiguous amino acids truncated or deleted from the amino acid sequence of SEQ ID NO:2,4, 6.
- 10 Fragments of the polypeptides of the invention may be employed for producing the corresponding full-length polypeptide by peptide synthesis; therefore, these fragments may be employed as intermediates for producing the full-length polypeptides of the invention.
- 15 Particularly preferred are variants in which several, 5-10, 1-5, 1-3, 1-2 or 1 amino acids are substituted, deleted, or added in any combination.

- The polypeptides, or immunogenic fragments, of the invention may be in the form of the "mature" protein or may be a part of a larger protein such as a precursor or a fusion protein. It is often advantageous to include an additional amino acid sequence which contains secretory or leader sequences, pro-sequences, sequences which aid in purification such as multiple histidine residues, or an additional sequence for stability during recombinant production. Furthermore, addition of exogenous polypeptide or lipid tail or polynucleotide sequences to increase the immunogenic potential of the final molecule is also considered.
- 20
- 25

In one aspect, the invention relates to genetically engineered soluble fusion proteins comprising a polypeptide of the present invention, or a fragment thereof, and various portions of the constant regions of heavy or light chains of immunoglobulins of various

subclasses (IgG, IgM, IgA, IgE). Preferred as an immunoglobulin is the constant part of the heavy chain of human IgG, particularly IgG1, where fusion takes place at the hinge region. In a particular embodiment, the Fc part can be removed simply by incorporation of a cleavage sequence which can be cleaved with blood clotting factor Xa.

5

Furthermore, this invention relates to processes for the preparation of these fusion proteins by genetic engineering, and to the use thereof for drug screening, diagnosis and therapy. A further aspect of the invention also relates to polynucleotides encoding such fusion proteins. Examples of fusion protein technology can be found in International Patent Application Nos. WO94/29458 and WO94/22914.

10

The proteins may be chemically conjugated, or expressed as recombinant fusion proteins allowing increased levels to be produced in an expression system as compared to non-fused protein. The fusion partner may assist in providing T helper epitopes (immunological fusion partner), preferably T helper epitopes recognised by humans, or assist in expressing the protein (expression enhancer) at higher yields than the native recombinant protein. Preferably the fusion partner will be both an immunological fusion partner and expression enhancing partner.

15

Fusion partners include protein D from *Haemophilus influenzae* and the non-structural protein from influenzae virus, NS1 (hemagglutinin). Another fusion partner is the protein known as LytA. Preferably the C terminal portion of the molecule is used. LytA is derived from *Streptococcus pneumoniae* which synthesize an N-acetyl-L-alanine amidase, amidase LytA, (coded by the lytA gene {Gene, 43 (1986) page 265-272}) an autolysin that specifically degrades certain bonds in the peptidoglycan backbone. The C-terminal domain of the LytA protein is responsible for the affinity to the choline or to some choline analogues such as DEAE. This property has been exploited for the development of E.coli C-LytA expressing plasmids useful for expression of fusion proteins. Purification of hybrid proteins containing the C-LytA

20
25

fragment at its amino terminus has been described {Biotechnology: 10, (1992) page 795-798}. It is possible to use the repeat portion of the LytA molecule found in the C terminal end starting at residue 178, for example residues 188 - 305.

- 5 The present invention also includes variants of the aforementioned polypeptides, that is polypeptides that vary from the referents by conservative amino acid substitutions, whereby a residue is substituted by another with like characteristics. Typical such substitutions are among Ala, Val, Leu and Ile; among Ser and Thr; among the acidic residues Asp and Glu; among Asn and Gln; and among the basic residues Lys and Arg; or
10 aromatic residues Phe and Tyr.

Polypeptides of the present invention can be prepared in any suitable manner. Such polypeptides include isolated naturally occurring polypeptides, recombinantly produced polypeptides, synthetically produced polypeptides, or polypeptides produced by a
15 combination of these methods. Means for preparing such polypeptides are well understood in the art.

It is most preferred that a polypeptide of the invention is derived from *Neisseria meningitidis*, however, it may preferably be obtained from other organisms of the same
20 taxonomic genus. A polypeptide of the invention may also be obtained, for example, from organisms of the same taxonomic family or order.

Polynucleotides

It is an object of the invention to provide polynucleotides that encode BASB013
25 polypeptides, particularly polynucleotides that encode the polypeptide herein designated BASB013.

In a particularly preferred embodiment of the invention the polynucleotide comprises a region encoding BASB013 polypeptides comprising a sequence set out in SEQ ID NO:1,3, 5 which includes a full length gene, or a variant thereof.

- 5 The BASB013 polynucleotides provided in SEQ ID NO:1,3, 5 are the BASB013 polynucleotides from *Neisseria meningitidis* strains ATCC13090 and H44/76.

As a further aspect of the invention there are provided isolated nucleic acid molecules encoding and/or expressing BASB013 polypeptides and polynucleotides, particularly

- 10 *Neisseria meningitidis* BASB013 polypeptides and polynucleotides, including, for example, unprocessed RNAs, ribozyme RNAs, mRNAs, cDNAs, genomic DNAs, B- and Z-DNAs. Further embodiments of the invention include biologically, diagnostically, prophylactically, clinically or therapeutically useful polynucleotides and polypeptides, and variants thereof, and compositions comprising the same.

15

Another aspect of the invention relates to isolated polynucleotides, including at least one full length gene, that encodes a BASB013 polypeptide having a deduced amino acid sequence of SEQ ID NO:2,4, 6 and polynucleotides closely related thereto and variants thereof.

- 20 In another particularly preferred embodiment of the invention there is a BASB013 polypeptide from *Neisseria meningitidis* comprising or consisting of an amino acid sequence of SEQ ID NO:2,4, 6 or a variant thereof.

- Using the information provided herein, such as a polynucleotide sequence set out in SEQ ID
25 NO:1, 3, 5, a polynucleotide of the invention encoding BASB013 polypeptide may be obtained using standard cloning and screening methods, such as those for cloning and sequencing chromosomal DNA fragments from bacteria using *Neisseria meningitidis* cells as starting material, followed by obtaining a full length clone. For example, to obtain a polynucleotide sequence of the invention, such as a polynucleotide sequence given in

SEQ ID NO:1,3,5, typically a library of clones of chromosomal DNA of *Neisseria meningitidis* in *E.coli* or some other suitable host is probed with a radiolabeled oligonucleotide, preferably a 17-mer or longer, derived from a partial sequence. Clones carrying DNA identical to that of the probe can then be distinguished using stringent hybridization conditions. By sequencing the individual clones thus identified by hybridization with sequencing primers designed from the original polypeptide or polynucleotide sequence it is then possible to extend the polynucleotide sequence in both directions to determine a full length gene sequence. Conveniently, such sequencing is performed, for example, using denatured double stranded DNA prepared from a plasmid clone. Suitable techniques are described by Maniatis, T., Fritsch, E.F. and Sambrook et al., *MOLECULAR CLONING, A LABORATORY MANUAL*, 2nd Ed.; Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York (1989). (see in particular Screening By Hybridization 1.90 and Sequencing Denatured Double-Stranded DNA Templates 13.70). Direct genomic DNA sequencing may also be performed to obtain a full length gene sequence. Illustrative of the invention, each polynucleotide set out in SEQ ID NO:1,3, 5 was discovered in a DNA library derived from *Neisseria meningitidis*.

Moreover, each DNA sequence set out in SEQ ID NO:1,3, 5 contains an open reading frame encoding a protein having about the number of amino acid residues set forth in SEQ ID NO:2,4, 6 with a deduced molecular weight that can be calculated using amino acid residue molecular weight values well known to those skilled in the art.

The polynucleotide of SEQ ID NO:1, between the start codon at nucleotide number 1 and the stop codon which begins at nucleotide number 1498 of SEQ ID NO:1, encodes the polypeptide of SEQ ID NO:2.

The polynucleotide of SEQ ID NO:3, between the start codon at nucleotide number 1 and the stop codon which begins at nucleotide number 1498 of SEQ ID NO:3, encodes the polypeptide of SEQ ID NO:4.

The polynucleotide of SEQ ID NO:5, between the start codon at nucleotide number 1 and the stop codon which begins at nucleotide number 1498 of SEQ ID NO:5, encodes the polypeptide of SEQ ID NO:6.

- 5 In a further aspect, the present invention provides for an isolated polynucleotide comprising or consisting of:
- (a) a polynucleotide sequence which has at least 85% identity, more preferably at least 90% identity, yet more preferably at least 95% identity, even more preferably at least 97-99% or exact identity to SEQ ID NO:1,3, 5 over the entire length of SEQ ID NO:1,3, 5 respectively; or
 - 10 (b) a polynucleotide sequence encoding a polypeptide which has at least 85% identity, more preferably at least 90% identity, yet more preferably at least 95% identity, even more preferably at least 97-99% or 100% exact, to the amino acid sequence of SEQ ID NO:2,4, 6 over the entire length of SEQ ID NO:2,4, 6 respectively.

15

A polynucleotide encoding a polypeptide of the present invention, including homologs and orthologs from species other than *Neisseria meningitidis*, may be obtained by a process which comprises the steps of screening an appropriate library under stringent hybridization conditions (for example, using a temperature in the range of 45 – 65°C and an SDS

20 concentration from 0.1 – 1%) with a labeled or detectable probe consisting of or comprising the sequence of SEQ ID NO: 1, 3, 5 or a fragment thereof; and isolating a full-length gene and/or genomic clones containing said polynucleotide sequence.

The invention provides a polynucleotide sequence identical over its entire length to a coding

25 sequence (open reading frame) in SEQ ID NO: 1, 3, 5. Also provided by the invention is a coding sequence for a mature polypeptide or a fragment thereof, by itself as well as a coding sequence for a mature polypeptide or a fragment in reading frame with another coding sequence, such as a sequence encoding a leader or secretory sequence, a pre-, or pro- or prepro-protein sequence. The polynucleotide of the invention may also contain at least one

non-coding sequence, including for example, but not limited to at least one non-coding 5' and 3' sequence, such as the transcribed but non-translated sequences, termination signals (such as rho-dependent and rho-independent termination signals), ribosome binding sites, Kozak sequences, sequences that stabilize mRNA, introns, and polyadenylation signals.

5 The polynucleotide sequence may also comprise additional coding sequence encoding additional amino acids. For example, a marker sequence that facilitates purification of the fused polypeptide can be encoded. In certain embodiments of the invention, the marker sequence is a hexa-histidine peptide, as provided in the pQE vector (Qiagen, Inc.) and described in Gentz *et al.*, *Proc. Natl. Acad. Sci., USA* 86: 821-824 (1989), or an HA peptide tag (Wilson *et al.*, *Cell* 37: 767 (1984), both of which may be useful in purifying
10 polypeptide sequence fused to them. Polynucleotides of the invention also include, but are not limited to, polynucleotides comprising a structural gene and its naturally associated sequences that control gene expression.

15 The nucleotide sequence encoding BASB013 polypeptide of SEQ ID NO:2,4, 6 may be identical to the polypeptide encoding sequence contained in nucleotides 1 to 1497 of SEQ ID NO:1, or the polypeptide encoding sequence contained in nucleotides 1 to 1497 of SEQ ID NO:3; or the polypeptide encoding sequence contained in nucleotides 1 to 1497 of SEQ ID NO:5, respectively. Alternatively it may be a sequence, which as a result of the
20 redundancy (degeneracy) of the genetic code, also encodes the polypeptide of SEQ ID NO:2,4, 6 .

The term "polynucleotide encoding a polypeptide" as used herein encompasses polynucleotides that include a sequence encoding a polypeptide of the invention,
25 particularly a bacterial polypeptide and more particularly a polypeptide of the *Neisseria meningitidis* BASB013 having an amino acid sequence set out in SEQ ID NO:2,4, 6 . The term also encompasses polynucleotides that include a single continuous region or discontinuous regions encoding the polypeptide (for example, polynucleotides interrupted by integrated phage, an integrated insertion sequence, an integrated vector sequence, an

integrated transposon sequence, or due to RNA editing or genomic DNA reorganization) together with additional regions, that also may contain coding and/or non-coding sequences.

The invention further relates to variants of the polynucleotides described herein that encode
5 variants of a polypeptide having a deduced amino acid sequence of SEQ ID NO:2,4, 6 .

Fragments of polynucleotides of the invention may be used, for example, to synthesize full-length polynucleotides of the invention.

Further particularly preferred embodiments are polynucleotides encoding BASB013
10 variants, that have the amino acid sequence of BASB013 polypeptide of SEQ ID NO:2,4, 6 in which several, a few, 5 to 10, 1 to 5, 1 to 3, 2, 1 or no amino acid residues are substituted, modified, deleted and/or added, in any combination. Especially preferred among these are silent substitutions, additions and deletions, that do not alter the properties and activities of BASB013 polypeptide.

15 Further preferred embodiments of the invention are polynucleotides that are at least 85% identical over their entire length to a polynucleotide encoding BASB013 polypeptide having an amino acid sequence set out in SEQ ID NO:2,4, 6 , and polynucleotides that are complementary to such polynucleotides. In this regard, polynucleotides at least 90%
20 identical over their entire length to the same are particularly preferred, and among these particularly preferred polynucleotides, those with at least 95% are especially preferred. Furthermore, those with at least 97% are highly preferred among those with at least 95%, and among these those with at least 98% and at least 99% are particularly highly preferred, with at least 99% being the more preferred.

25 Preferred embodiments are polynucleotides encoding polypeptides that retain substantially the same biological function or activity as the mature polypeptide encoded by a DNA of SEQ ID NO:1, 3, 5.

In accordance with certain preferred embodiments of this invention there are provided polynucleotides that hybridize, particularly under stringent conditions, to BASB013 polynucleotide sequences, such as those polynucleotides in SEQ ID NO:1, 3, 5.

- 5 The invention further relates to polynucleotides that hybridize to the polynucleotide sequences provided herein. In this regard, the invention especially relates to polynucleotides that hybridize under stringent conditions to the polynucleotides described herein. As herein used, the terms "stringent conditions" and "stringent hybridization conditions" mean hybridization occurring only if there is at least 95% and preferably at least 97% identity
- 10 between the sequences. A specific example of stringent hybridization conditions is overnight incubation at 42°C in a solution comprising: 50% formamide, 5x SSC (150mM NaCl, 15mM trisodium citrate), 50 mM sodium phosphate (pH7.6), 5x Denhardt's solution, 10% dextran sulfate, and 20 micrograms/ml of denatured, sheared salmon sperm DNA, followed by washing the hybridization support in 0.1x SSC at about 65°C.
- 15 Hybridization and wash conditions are well known and exemplified in Sambrook, *et al.*, Molecular Cloning: A Laboratory Manual, Second Edition, Cold Spring Harbor, N.Y., (1989), particularly Chapter 11 therein. Solution hybridization may also be used with the polynucleotide sequences provided by the invention.
- 20 The invention also provides a polynucleotide consisting of or comprising a polynucleotide sequence obtained by screening an appropriate library containing the complete gene for a polynucleotide sequence set forth in SEQ ID NO:1, 3, 5 under stringent hybridization conditions with a probe having the sequence of said polynucleotide sequence set forth in SEQ ID NO:1, 3, 5 or a fragment thereof; and isolating said polynucleotide sequence.
- 25 Fragments useful for obtaining such a polynucleotide include, for example, probes and primers fully described elsewhere herein.

As discussed elsewhere herein regarding polynucleotide assays of the invention, for instance, the polynucleotides of the invention, may be used as a hybridization probe for

RNA, cDNA and genomic DNA to isolate full-length cDNAs and genomic clones encoding BASB013 and to isolate cDNA and genomic clones of other genes that have a high identity, particularly high sequence identity, to the BASB013 gene. Such probes generally will comprise at least 15 nucleotide residues or base pairs. Preferably, such probes will have at least 30 nucleotide residues or base pairs and may have at least 50 nucleotide residues or base pairs. Particularly preferred probes will have at least 20 nucleotide residues or base pairs and will have less than 30 nucleotide residues or base pairs.

A coding region of a BASB013 gene may be isolated by screening using a DNA sequence provided in SEQ ID NO:1, 3, 5 to synthesize an oligonucleotide probe. A labeled oligonucleotide having a sequence complementary to that of a gene of the invention is then used to screen a library of cDNA, genomic DNA or mRNA to determine which members of the library the probe hybridizes to.

There are several methods available and well known to those skilled in the art to obtain full-length DNAs, or extend short DNAs, for example those based on the method of Rapid Amplification of cDNA ends (RACE) (see, for example, Frohman, *et al.*, *PNAS USA* 85: 8998-9002, 1988). Recent modifications of the technique, exemplified by the Marathon™ technology (Clontech Laboratories Inc.) for example, have significantly simplified the search for longer cDNAs. In the Marathon™ technology, cDNAs have been prepared from mRNA extracted from a chosen tissue and an 'adaptor' sequence ligated onto each end. Nucleic acid amplification (PCR) is then carried out to amplify the "missing" 5' end of the DNA using a combination of gene specific and adaptor specific oligonucleotide primers. The PCR reaction is then repeated using "nested" primers, that is, primers designed to anneal within the amplified product (typically an adaptor specific primer that anneals further 3' in the adaptor sequence and a gene specific primer that anneals further 5' in the selected gene sequence). The products of this reaction can then be analyzed by DNA sequencing and a full-length DNA constructed either by joining the product directly

to the existing DNA to give a complete sequence, or carrying out a separate full-length PCR using the new sequence information for the design of the 5' primer.

5 The polynucleotides and polypeptides of the invention may be employed, for example, as research reagents and materials for discovery of treatments of and diagnostics for diseases, particularly human diseases, as further discussed herein relating to polynucleotide assays.

10 The polynucleotides of the invention that are oligonucleotides derived from a sequence of SEQ ID NOS:1 – 6 may be used in the processes herein as described, but preferably for PCR, to determine whether or not the polynucleotides identified herein in whole or in part are transcribed in bacteria in infected tissue. It is recognized that such sequences will also have utility in diagnosis of the stage of infection and type of infection the pathogen has attained.

15 The invention also provides polynucleotides that encode a polypeptide that is the mature protein plus additional amino or carboxyl-terminal amino acids, or amino acids interior to the mature polypeptide (when the mature form has more than one polypeptide chain, for instance). Such sequences may play a role in processing of a protein from precursor to a mature form, may allow protein transport, may lengthen or shorten protein half-life or may
20 facilitate manipulation of a protein for assay or production, among other things. As generally is the case *in vivo*, the additional amino acids may be processed away from the mature protein by cellular enzymes.

25 For each and every polynucleotide of the invention there is provided a polynucleotide complementary to it. It is preferred that these complementary polynucleotides are fully complementary to each polynucleotide with which they are complementary.

A precursor protein, having a mature form of the polypeptide fused to one or more prosequences may be an inactive form of the polypeptide. When prosequences are removed

such inactive precursors generally are activated. Some or all of the prosequences may be removed before activation. Generally, such precursors are called proproteins.

In addition to the standard A, G, C, T/U representations for nucleotides, the term "N" may also be used in describing certain polynucleotides of the invention. "N" means that any of the four DNA or RNA nucleotides may appear at such a designated position in the DNA or RNA sequence, except it is preferred that N is not a nucleic acid that when taken in combination with adjacent nucleotide positions, when read in the correct reading frame, would have the effect of generating a premature termination codon in such reading frame.

In sum, a polynucleotide of the invention may encode a mature protein, a mature protein plus a leader sequence (which may be referred to as a preprotein), a precursor of a mature protein having one or more prosequences that are not the leader sequences of a preprotein, or a preproprotein, which is a precursor to a proprotein, having a leader sequence and one or more prosequences, which generally are removed during processing steps that produce active and mature forms of the polypeptide.

In accordance with an aspect of the invention, there is provided the use of a polynucleotide of the invention for therapeutic or prophylactic purposes, in particular genetic immunization.

The use of a polynucleotide of the invention in genetic immunization will preferably employ a suitable delivery method such as direct injection of plasmid DNA into muscles (Wolff *et al.*, *Hum Mol Genet* (1992) 1: 363, Manthorpe *et al.*, *Hum. Gene Ther.* (1983) 4: 419), delivery of DNA complexed with specific protein carriers (Wu *et al.*, *J Biol Chem.* (1989) 264: 16985), coprecipitation of DNA with calcium phosphate (Benvenisty & Reshef, *PNAS USA*, (1986) 83: 9551), encapsulation of DNA in various forms of liposomes (Kaneda *et al.*, *Science* (1989) 243: 375), particle bombardment (Tang *et al.*,

Nature (1992) 356:152, Eisenbraun *et al.*, *DNA Cell Biol* (1993) 12: 791) and *in vivo* infection using cloned retroviral vectors (Seeger *et al.*, *PNAS USA* (1984) 81: 5849).

5 Vectors, Host Cells, Expression Systems

The invention also relates to vectors that comprise a polynucleotide or polynucleotides of the invention, host cells that are genetically engineered with vectors of the invention and the production of polypeptides of the invention by recombinant techniques. Cell-free translation systems can also be employed to produce such proteins using RNAs derived
10 from the DNA constructs of the invention.

Recombinant polypeptides of the present invention may be prepared by processes well known in those skilled in the art from genetically engineered host cells comprising expression systems. Accordingly, in a further aspect, the present invention relates to
15 expression systems that comprise a polynucleotide or polynucleotides of the present invention, to host cells which are genetically engineered with such expression systems, and to the production of polypeptides of the invention by recombinant techniques.

For recombinant production of the polypeptides of the invention, host cells can be
20 genetically engineered to incorporate expression systems or portions thereof or polynucleotides of the invention. Introduction of a polynucleotide into the host cell can be effected by methods described in many standard laboratory manuals, such as Davis, *et al.*, *BASIC METHODS IN MOLECULAR BIOLOGY*, (1986) and Sambrook, *et al.*, *MOLECULAR CLONING: A LABORATORY MANUAL*, 2nd Ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y. (1989), such as, calcium phosphate
25 transfection, DEAE-dextran mediated transfection, transvection, microinjection, cationic lipid-mediated transfection, electroporation, transduction, scrape loading, ballistic introduction and infection.

Representative examples of appropriate hosts include bacterial cells, such as cells of streptococci, staphylococci, enterococci, *E. coli*, streptomyces, cyanobacteria, *Bacillus subtilis*, *Moraxella catarrhalis*, *Haemophilus influenzae* and *Neisseria meningitidis*; fungal cells, such as cells of a yeast, *Kluveromyces*, *Saccharomyces*, a basidiomycete, *Candida albicans* and *Aspergillus*; insect cells such as cells of *Drosophila* S2 and *Spodoptera* Sf9; animal cells such as CHO, COS, HeLa, C127, 3T3, BHK, 293, CV-1 and Bowes melanoma cells; and plant cells, such as cells of a gymnosperm or angiosperm.

A great variety of expression systems can be used to produce the polypeptides of the invention. Such vectors include, among others, chromosomal-, episomal- and virus-derived vectors, for example, vectors derived from bacterial plasmids, from bacteriophage, from transposons, from yeast episomes, from insertion elements, from yeast chromosomal elements, from viruses such as baculoviruses, papova viruses, such as SV40, vaccinia viruses, adenoviruses, fowl pox viruses, pseudorabies viruses, picornaviruses, retroviruses, and alphaviruses and vectors derived from combinations thereof, such as those derived from plasmid and bacteriophage genetic elements, such as cosmids and phagemids. The expression system constructs may contain control regions that regulate as well as engender expression. Generally, any system or vector suitable to maintain, propagate or express polynucleotides and/or to express a polypeptide in a host may be used for expression in this regard. The appropriate DNA sequence may be inserted into the expression system by any of a variety of well-known and routine techniques, such as, for example, those set forth in Sambrook *et al.*, *MOLECULAR CLONING, A LABORATORY MANUAL*, (*supra*).

In recombinant expression systems in eukaryotes, for secretion of a translated protein into the lumen of the endoplasmic reticulum, into the periplasmic space or into the extracellular environment, appropriate secretion signals may be incorporated into the expressed polypeptide. These signals may be endogenous to the polypeptide or they may be heterologous signals.

Polypeptides of the present invention can be recovered and purified from recombinant cell cultures by well-known methods including ammonium sulfate or ethanol precipitation, acid extraction, anion or cation exchange chromatography, phosphocellulose chromatography, hydrophobic interaction chromatography, affinity chromatography, hydroxylapatite chromatography and lectin chromatography. Most preferably, ion metal affinity chromatography (IMAC) is employed for purification. Well known techniques for refolding proteins may be employed to regenerate active conformation when the polypeptide is denatured during intracellular synthesis, isolation and or purification.

The expression system may also be a recombinant live microorganism, such as a virus or bacterium. The gene of interest can be inserted into the genome of a live recombinant virus or bacterium. Inoculation and *in vivo* infection with this live vector will lead to *in vivo* expression of the antigen and induction of immune responses. Viruses and bacteria used for this purpose are for instance: poxviruses (e.g; vaccinia, fowlpox, canarypox), alphaviruses (Sindbis virus, Semliki Forest Virus, Venezuelan Equine Encephalitis Virus), adenoviruses, adeno-associated virus, picornaviruses (poliovirus, rhinovirus), herpesviruses (varicella zoster virus, etc), Listeria, Salmonella, Shigella, Neisseria, BCG. These viruses and bacteria can be virulent, or attenuated in various ways in order to obtain live vaccines. Such live vaccines also form part of the invention.

Diagnostic, Prognostic, Serotyping and Mutation Assays

This invention is also related to the use of BASB013 polynucleotides and polypeptides of the invention for use as diagnostic reagents. Detection of BASB013 polynucleotides and/or polypeptides in a eukaryote, particularly a mammal, and especially a human, will provide a diagnostic method for diagnosis of disease, staging of disease or response of an infectious organism to drugs. Eukaryotes, particularly mammals, and especially humans, particularly those infected or suspected to be infected with an organism comprising the BASB013 gene or protein, may be detected at the nucleic acid or amino acid level by a variety of well known techniques as well as by methods provided herein.

Polypeptides and polynucleotides for prognosis, diagnosis or other analysis may be obtained from a putatively infected and/or infected individual's bodily materials. Polynucleotides from any of these sources, particularly DNA or RNA, may be used directly for detection or may be amplified enzymatically by using PCR or any other amplification technique prior to analysis. RNA, particularly mRNA, cDNA and genomic DNA may also be used in the same ways. Using amplification, characterization of the species and strain of infectious or resident organism present in an individual, may be made by an analysis of the genotype of a selected polynucleotide of the organism. Deletions and insertions can be detected by a change in size of the amplified product in comparison to a genotype of a reference sequence selected from a related organism, preferably a different species of the same genus or a different strain of the same species. Point mutations can be identified by hybridizing amplified DNA to labeled BASB013 polynucleotide sequences. Perfectly or significantly matched sequences can be distinguished from imperfectly or more significantly mismatched duplexes by DNase or RNase digestion, for DNA or RNA respectively, or by detecting differences in melting temperatures or renaturation kinetics. Polynucleotide sequence differences may also be detected by alterations in the electrophoretic mobility of polynucleotide fragments in gels as compared to a reference sequence. This may be carried out with or without denaturing agents. Polynucleotide differences may also be detected by direct DNA or RNA sequencing. See, for example, Myers *et al.*, *Science*, 230: 1242 (1985). Sequence changes at specific locations also may be revealed by nuclease protection assays, such as RNase, V1 and S1 protection assay or a chemical cleavage method. See, for example, Cotton *et al.*, *Proc. Natl. Acad. Sci., USA*, 85: 4397-4401 (1985).

In another embodiment, an array of oligonucleotides probes comprising BASB013 nucleotide sequence or fragments thereof can be constructed to conduct efficient screening of, for example, genetic mutations, serotype, taxonomic classification or identification. Array technology methods are well known and have general applicability and can be used to

address a variety of questions in molecular genetics including gene expression, genetic linkage, and genetic variability (see, for example, Chee *et al.*, *Science*, 274: 610 (1996)).

Thus in another aspect, the present invention relates to a diagnostic kit which comprises:

- 5 (a) a polynucleotide of the present invention, preferably the nucleotide sequence of SEQ ID NO:1, 3, 5 or a fragment thereof ;
- (b) a nucleotide sequence complementary to that of (a);
- (c) a polypeptide of the present invention, preferably the polypeptide of SEQ ID NO:2,4, 6 or a fragment thereof; or
- 10 (d) an antibody to a polypeptide of the present invention, preferably to the polypeptide of SEQ ID NO:2,4, 6 .

It will be appreciated that in any such kit, (a), (b), (c) or (d) may comprise a substantial component. Such a kit will be of use in diagnosing a disease or susceptibility to a disease,
15 among others.

This invention also relates to the use of polynucleotides of the present invention as diagnostic reagents. Detection of a mutated form of a polynucleotide of the invention, preferable, SEQ ID NO:1, 3, 5 which is associated with a disease or pathogenicity will
20 provide a diagnostic tool that can add to, or define, a diagnosis of a disease, a prognosis of a course of disease, a determination of a stage of disease, or a susceptibility to a disease, which results from under-expression, over-expression or altered expression of the polynucleotide. Organisms, particularly infectious organisms, carrying mutations in such polynucleotide may be detected at the polynucleotide level by a variety of techniques, such
25 as those described elsewhere herein.

Cells from an organism carrying mutations or polymorphisms (allelic variations) in a polynucleotide and/or polypeptide of the invention may also be detected at the polynucleotide or polypeptide level by a variety of techniques, to allow for serotyping, for

example. For example, RT-PCR can be used to detect mutations in the RNA. It is particularly preferred to use RT-PCR in conjunction with automated detection systems, such as, for example, GeneScan. RNA, cDNA or genomic DNA may also be used for the same purpose, PCR. As an example, PCR primers complementary to a polynucleotide encoding BASB013 polypeptide can be used to identify and analyze mutations.

The invention further provides primers with 1, 2, 3 or 4 nucleotides removed from the 5' and/or the 3' end. These primers may be used for, among other things, amplifying BASB013 DNA and/or RNA isolated from a sample derived from an individual, such as a bodily material. The primers may be used to amplify a polynucleotide isolated from an infected individual, such that the polynucleotide may then be subject to various techniques for elucidation of the polynucleotide sequence. In this way, mutations in the polynucleotide sequence may be detected and used to diagnose and/or prognose the infection or its stage or course, or to serotype and/or classify the infectious agent.

The invention further provides a process for diagnosing disease, preferably bacterial infections, more preferably infections caused by *Neisseria meningitidis*, comprising determining from a sample derived from an individual, such as a bodily material, an increased level of expression of polynucleotide having a sequence of SEQ ID NO:1, 3, 5. Increased or decreased expression of a BASB013 polynucleotide can be measured using any one of the methods well known in the art for the quantitation of polynucleotides, such as, for example, amplification, PCR, RT-PCR, RNase protection, Northern blotting, spectrometry and other hybridization methods.

In addition, a diagnostic assay in accordance with the invention for detecting over-expression of BASB013 polypeptide compared to normal control tissue samples may be used to detect the presence of an infection, for example. Assay techniques that can be used to determine levels of a BASB013 polypeptide, in a sample derived from a host, such as a bodily material, are well-known to those of skill in the art. Such assay methods include

radioimmunoassays, competitive-binding assays, Western Blot analysis, antibody sandwich assays, antibody detection and ELISA assays.

The polynucleotides of the invention may be used as components of polynucleotide arrays, preferably high density arrays or grids. These high density arrays are particularly useful for diagnostic and prognostic purposes. For example, a set of spots each comprising a different gene, and further comprising a polynucleotide or polynucleotides of the invention, may be used for probing, such as using hybridization or nucleic acid amplification, using a probe obtained or derived from a bodily sample, to determine the presence of a particular polynucleotide sequence or related sequence in an individual. Such a presence may indicate the presence of a pathogen, particularly *Neisseria meningitidis*, and may be useful in diagnosing and/or prognosing disease or a course of disease. A grid comprising a number of variants of the polynucleotide sequence of SEQ ID NO:1, 3, 5 are preferred. Also preferred is a grid comprising a number of variants of a polynucleotide sequence encoding the polypeptide sequence of SEQ ID NO:2,4, 6 .

Antibodies

The polypeptides and polynucleotides of the invention or variants thereof, or cells expressing the same can be used as immunogens to produce antibodies immunospecific for such polypeptides or polynucleotides respectively.

In certain preferred embodiments of the invention there are provided antibodies against BASB013 polypeptides or polynucleotides.

Antibodies generated against the polypeptides or polynucleotides of the invention can be obtained by administering the polypeptides and/or polynucleotides of the invention, or epitope-bearing fragments of either or both, analogues of either or both, or cells expressing either or both, to an animal, preferably a nonhuman, using routine protocols. For

preparation of monoclonal antibodies, any technique known in the art that provides antibodies produced by continuous cell line cultures can be used. Examples include various techniques, such as those in Kohler, G. and Milstein, C., *Nature* 256: 495-497 (1975); Kozbor *et al.*, *Immunology Today* 4: 72 (1983); Cole *et al.*, pg. 77-96 in *MONOCLONAL*
5 *ANTIBODIES AND CANCER THERAPY*, Alan R. Liss, Inc. (1985).

Techniques for the production of single chain antibodies (U.S. Patent No. 4,946,778) can be adapted to produce single chain antibodies to polypeptides or polynucleotides of this invention. Also, transgenic mice, or other organisms or animals, such as other mammals,
10 may be used to express humanized antibodies immunospecific to the polypeptides or polynucleotides of the invention.

Alternatively, phage display technology may be utilized to select antibody genes with binding activities towards a polypeptide of the invention either from repertoires of PCR
15 amplified v-genes of lymphocytes from humans screened for possessing anti-BASB013 or from naive libraries (McCafferty, *et al.*, (1990), *Nature* 348, 552-554; Marks, *et al.*, (1992) *Biotechnology* 10, 779-783). The affinity of these antibodies can also be improved by, for example, chain shuffling (Clackson *et al.*, (1991) *Nature* 352: 628).

20 The above-described antibodies may be employed to isolate or to identify clones expressing the polypeptides or polynucleotides of the invention to purify the polypeptides or polynucleotides by, for example, affinity chromatography.

Thus, among others, antibodies against BASB013-polypeptide or BASB013-polynucleotide
25 may be employed to treat infections, particularly bacterial infections.

Polypeptide variants include antigenically, epitopically or immunologically equivalent variants form a particular aspect of this invention.

Preferably, the antibody or variant thereof is modified to make it less immunogenic in the individual. For example, if the individual is human the antibody may most preferably be "humanized," where the complementarity determining region or regions of the hybridoma-derived antibody has been transplanted into a human monoclonal antibody, for example as described in Jones *et al.* (1986), *Nature* 321, 522-525 or Tempest *et al.*, (1991) *Biotechnology* 9, 266-273.

Antagonists and Agonists - Assays and Molecules

Polypeptides and polynucleotides of the invention may also be used to assess the binding of small molecule substrates and ligands in, for example, cells, cell-free preparations, chemical libraries, and natural product mixtures. These substrates and ligands may be natural substrates and ligands or may be structural or functional mimetics. See, *e.g.*, Coligan *et al.*, *Current Protocols in Immunology* 1(2): Chapter 5 (1991).

The screening methods may simply measure the binding of a candidate compound to the polypeptide or polynucleotide, or to cells or membranes bearing the polypeptide or polynucleotide, or a fusion protein of the polypeptide by means of a label directly or indirectly associated with the candidate compound. Alternatively, the screening method may involve competition with a labeled competitor. Further, these screening methods may test whether the candidate compound results in a signal generated by activation or inhibition of the polypeptide or polynucleotide, using detection systems appropriate to the cells comprising the polypeptide or polynucleotide. Inhibitors of activation are generally assayed in the presence of a known agonist and the effect on activation by the agonist by the presence of the candidate compound is observed. Constitutively active polypeptide and/or constitutively expressed polypeptides and polynucleotides may be employed in screening methods for inverse agonists or inhibitors, in the absence of an agonist or inhibitor, by testing whether the candidate compound results in inhibition of activation of the polypeptide or polynucleotide, as the case may be. Further, the screening methods

may simply comprise the steps of mixing a candidate compound with a solution containing a polypeptide or polynucleotide of the present invention, to form a mixture, measuring BASB013 polypeptide and/or polynucleotide activity in the mixture, and comparing the BASB013 polypeptide and/or polynucleotide activity of the mixture to a standard. Fusion proteins, such as those made from Fc portion and BASB013 polypeptide, as hereinbefore described, can also be used for high-throughput screening assays to identify antagonists of the polypeptide of the present invention, as well as of phylogenetically and and/or functionally related polypeptides (see D. Bennett *et al.*, J Mol Recognition, 8:52-58 (1995); and K. Johanson *et al.*, J Biol Chem, 270(16):9459-9471 (1995)).

The polynucleotides, polypeptides and antibodies that bind to and/or interact with a polypeptide of the present invention may also be used to configure screening methods for detecting the effect of added compounds on the production of mRNA and/or polypeptide in cells. For example, an ELISA assay may be constructed for measuring secreted or cell associated levels of polypeptide using monoclonal and polyclonal antibodies by standard methods known in the art. This can be used to discover agents which may inhibit or enhance the production of polypeptide (also called antagonist or agonist, respectively) from suitably manipulated cells or tissues.

The invention also provides a method of screening compounds to identify those which enhance (agonist) or block (antagonist) the action of BASB013 polypeptides or polynucleotides, particularly those compounds that are bacteristatic and/or bactericidal. The method of screening may involve high-throughput techniques. For example, to screen for agonists or antagonists, a synthetic reaction mix, a cellular compartment, such as a membrane, cell envelope or cell wall, or a preparation of any thereof, comprising BASB013 polypeptide and a labeled substrate or ligand of such polypeptide is incubated in the absence or the presence of a candidate molecule that may be a BASB013 agonist or antagonist. The ability of the candidate molecule to agonize or antagonize the BASB013 polypeptide is

reflected in decreased binding of the labeled ligand or decreased production of product from such substrate. Molecules that bind gratuitously, *i.e.*, without inducing the effects of BASB013 polypeptide are most likely to be good antagonists. Molecules that bind well and, as the case may be, increase the rate of product production from substrate, increase signal transduction, or increase chemical channel activity are agonists. Detection of the rate or level of, as the case may be, production of product from substrate, signal transduction, or chemical channel activity may be enhanced by using a reporter system. Reporter systems that may be useful in this regard include but are not limited to colorimetric, labeled substrate converted into product, a reporter gene that is responsive to changes in BASB013 polynucleotide or polypeptide activity, and binding assays known in the art.

Another example of an assay for BASB013 agonists is a competitive assay that combines BASB013 and a potential agonist with BASB013-binding molecules, recombinant BASB013 binding molecules, natural substrates or ligands, or substrate or ligand mimetics, under appropriate conditions for a competitive inhibition assay. BASB013 can be labeled, such as by radioactivity or a colorimetric compound, such that the number of BASB013 molecules bound to a binding molecule or converted to product can be determined accurately to assess the effectiveness of the potential antagonist.

Potential antagonists include, among others, small organic molecules, peptides, polypeptides and antibodies that bind to a polynucleotide and/or polypeptide of the invention and thereby inhibit or extinguish its activity or expression. Potential antagonists also may be small organic molecules, a peptide, a polypeptide such as a closely related protein or antibody that binds the same sites on a binding molecule, such as a binding molecule, without inducing BASB013-induced activities, thereby preventing the action or expression of BASB013 polypeptides and/or polynucleotides by excluding BASB013 polypeptides and/or polynucleotides from binding.

Potential antagonists include a small molecule that binds to and occupies the binding site of the polypeptide thereby preventing binding to cellular binding molecules, such that normal biological activity is prevented. Examples of small molecules include but are not limited to small organic molecules, peptides or peptide-like molecules. Other potential antagonists include antisense molecules (see Okano, *J. Neurochem.* 56: 560 (1991); *OLIGODEOXYNUCLEOTIDES AS ANTISENSE INHIBITORS OF GENE EXPRESSION*, CRC Press, Boca Raton, FL (1988), for a description of these molecules). Preferred potential antagonists include compounds related to and variants of BASB006.

In a further aspect, the present invention relates to genetically engineered soluble fusion proteins comprising a polypeptide of the present invention, or a fragment thereof, and various portions of the constant regions of heavy or light chains of immunoglobulins of various subclasses (IgG, IgM, IgA, IgE). Preferred as an immunoglobulin is the constant part of the heavy chain of human IgG, particularly IgG1, where fusion takes place at the hinge region. In a particular embodiment, the Fc part can be removed simply by incorporation of a cleavage sequence which can be cleaved with blood clotting factor Xa. Furthermore, this invention relates to processes for the preparation of these fusion proteins by genetic engineering, and to the use thereof for drug screening, diagnosis and therapy. A further aspect of the invention also relates to polynucleotides encoding such fusion proteins. Examples of fusion protein technology can be found in International Patent Application Nos. WO94/29458 and WO94/22914.

Each of the polynucleotide sequences provided herein may be used in the discovery and development of antibacterial compounds. The encoded protein, upon expression, can be used as a target for the screening of antibacterial drugs. Additionally, the polynucleotide sequences encoding the amino terminal regions of the encoded protein or Shine-Delgarno or other translation facilitating sequences of the respective mRNA can be used to construct antisense sequences to control the expression of the coding sequence of interest.

The invention also provides the use of the polypeptide, polynucleotide, agonist or antagonist of the invention to interfere with the initial physical interaction between a pathogen or pathogens and a eukaryotic, preferably mammalian, host responsible for sequelae of infection. In particular, the molecules of the invention may be used: in the
5 prevention of adhesion of bacteria, in particular gram positive and/or gram negative bacteria, to eukaryotic, preferably mammalian, extracellular matrix proteins on in-dwelling devices or to extracellular matrix proteins in wounds; to block bacterial adhesion between eukaryotic, preferably mammalian, extracellular matrix proteins and bacterial
10 BASB013 proteins that mediate tissue damage and/or; to block the normal progression of pathogenesis in infections initiated other than by the implantation of in-dwelling devices or by other surgical techniques.

In accordance with yet another aspect of the invention, there are provided BASB013 agonists and antagonists, preferably bacteristatic or bactericidal agonists and antagonists.
15

The antagonists and agonists of the invention may be employed, for instance, to prevent, inhibit and/or treat diseases.

In a further aspect, the present invention relates to mimotopes of the polypeptide of the invention. A mimotope is a peptide sequence, sufficiently similar to the native peptide
20 (sequentially or structurally), which is capable of being recognised by antibodies which recognise the native peptide; or is capable of raising antibodies which recognise the native peptide when coupled to a suitable carrier.

25 Peptide mimotopes may be designed for a particular purpose by addition, deletion or substitution of elected amino acids. Thus, the peptides may be modified for the purposes of ease of conjugation to a protein carrier. For example, it may be desirable for some chemical conjugation methods to include a terminal cysteine. In addition it may be desirable for peptides conjugated to a protein carrier to include a hydrophobic terminus

distal from the conjugated terminus of the peptide, such that the free unconjugated end of the peptide remains associated with the surface of the carrier protein. Thereby presenting the peptide in a conformation which most closely resembles that of the peptide as found in the context of the whole native molecule. For example, the peptides
5 may be altered to have an N-terminal cysteine and a C-terminal hydrophobic amidated tail. Alternatively, the addition or substitution of a D-stereoisomer form of one or more of the amino acids may be performed to create a beneficial derivative, for example to enhance stability of the peptide.

10 Alternatively, peptide mimotopes may be identified using antibodies which are capable themselves of binding to the polypeptides of the present invention using techniques such as phage display technology (EP 0 552 267 B1). This technique, generates a large number of peptide sequences which mimic the structure of the native peptides and are, therefore, capable of binding to anti-native peptide antibodies, but may not necessarily themselves
15 share significant sequence homology to the native polypeptide.

Vaccines

Another aspect of the invention relates to a method for inducing an immunological response in an individual, particularly a mammal, preferably humans, which comprises
20 inoculating the individual with BASB013 polynucleotide and/or polypeptide, or a fragment or variant thereof, adequate to produce antibody and/ or T cell immune response to protect said individual from infection, particularly bacterial infection and most particularly *Neisseria meningitidis* infection. Also provided are methods whereby such immunological response slows bacterial replication. Yet another aspect of the invention
25 relates to a method of inducing immunological response in an individual which comprises delivering to such individual a nucleic acid vector, sequence or ribozyme to direct expression of BASB013 polynucleotide and/or polypeptide, or a fragment or a variant thereof, for expressing BASB013 polynucleotide and/or polypeptide, or a fragment or a variant thereof *in vivo* in order to induce an immunological response, such as, to produce

antibody and/ or T cell immune response. including, for example, cytokine-producing T cells or cytotoxic T cells, to protect said individual, preferably a human, from disease, whether that disease is already established within the individual or not. One example of administering the gene is by accelerating it into the desired cells as a coating on particles
5 or otherwise. Such nucleic acid vector may comprise DNA, RNA, a ribozyme, a modified nucleic acid, a DNA/RNA hybrid, a DNA-protein complex or an RNA-protein complex.

A further aspect of the invention relates to an immunological composition that when introduced into an individual, preferably a human, capable of having induced within it an
10 immunological response, induces an immunological response in such individual to a BASB013 polynucleotide and/or polypeptide encoded therefrom, wherein the composition comprises a recombinant BASB013 polynucleotide and/or polypeptide encoded therefrom and/or comprises DNA and/or RNA which encodes and expresses an antigen of said
BASB013 polynucleotide, polypeptide encoded therefrom, or other polypeptide of the
15 invention. The immunological response may be used therapeutically or prophylactically and may take the form of antibody immunity and/or cellular immunity, such as cellular immunity arising from CTL or CD4+ T cells.

A BASB013 polypeptide or a fragment thereof may be fused with co-protein or chemical
20 moiety which may or may not by itself produce antibodies, but which is capable of stabilizing the first protein and producing a fused or modified protein which will have antigenic and/or immunogenic properties, and preferably protective properties. Thus fused recombinant protein, preferably further comprises an antigenic co-protein, such as lipoprotein D from *Haemophilus influenzae*, Glutathione-S-transferase (GST) or beta-
25 galactosidase, or any other relatively large co-protein which solubilizes the protein and facilitates production and purification thereof. Moreover, the co-protein may act as an adjuvant in the sense of providing a generalized stimulation of the immune system of the organism receiving the protein. The co-protein may be attached to either the amino- or carboxy-terminus of the first protein.

Provided by this invention are compositions, particularly vaccine compositions, and methods comprising the polypeptides and/or polynucleotides of the invention and immunostimulatory DNA sequences, such as those described in Sato, Y. *et al.* Science
5 273: 352 (1996).

Also, provided by this invention are methods using the described polynucleotide or particular fragments thereof, which have been shown to encode non-variable regions of bacterial cell surface proteins, in polynucleotide constructs used in such genetic
10 immunization experiments in animal models of infection with *Neisseria meningitidis*. Such experiments will be particularly useful for identifying protein epitopes able to provoke a prophylactic or therapeutic immune response. It is believed that this approach will allow for the subsequent preparation of monoclonal antibodies of particular value, derived from the requisite organ of the animal successfully resisting or clearing infection,
15 for the development of prophylactic agents or therapeutic treatments of bacterial infection, particularly *Neisseria meningitidis* infection, in mammals, particularly humans.

The invention also includes a vaccine formulation which comprises an immunogenic recombinant polypeptide and/or polynucleotide of the invention together with a suitable
20 carrier, such as a pharmaceutically acceptable carrier. Since the polypeptides and polynucleotides may be broken down in the stomach, each is preferably administered parenterally, including, for example, administration that is subcutaneous, intramuscular, intravenous, or intradermal. Formulations suitable for parenteral administration include aqueous and non-aqueous sterile injection solutions which may contain anti-oxidants,
25 buffers, bacteristatic compounds and solutes which render the formulation isotonic with the bodily fluid, preferably the blood, of the individual; and aqueous and non-aqueous sterile suspensions which may include suspending agents or thickening agents. The formulations may be presented in unit-dose or multi-dose containers, for example, sealed

ampoules and vials and may be stored in a freeze-dried condition requiring only the addition of the sterile liquid carrier immediately prior to use.

5 The vaccine formulation of the invention may also include adjuvant systems for enhancing the immunogenicity of the formulation. Preferably the adjuvant system raises preferentially a TH1 type of response.

10 An immune response may be broadly distinguished into two extreme categories, being a humoral or cell mediated immune responses (traditionally characterised by antibody and cellular effector mechanisms of protection respectively). These categories of response have been termed TH1-type responses (cell-mediated response), and TH2-type immune responses (humoral response).

15 Extreme TH1-type immune responses may be characterised by the generation of antigen specific, haplotype restricted cytotoxic T lymphocytes, and natural killer cell responses. In mice TH1-type responses are often characterised by the generation of antibodies of the IgG2a subtype, whilst in the human these correspond to IgG1 type antibodies. TH2-type immune responses are characterised by the generation of a broad range of immunoglobulin isotypes including in mice IgG1, IgA, and IgM.

20 It can be considered that the driving force behind the development of these two types of immune responses are cytokines. High levels of TH1-type cytokines tend to favour the induction of cell mediated immune responses to the given antigen, whilst high levels of TH2-type cytokines tend to favour the induction of humoral immune responses to the
25 antigen.

The distinction of TH1 and TH2-type immune responses is not absolute. In reality an individual will support an immune response which is described as being predominantly TH1 or predominantly TH2. However, it is often convenient to consider the families of

cytokines in terms of that described in murine CD4 +ve T cell clones by Mosmann and Coffman (*Mosmann, T.R. and Coffman, R.L. (1989) TH1 and TH2 cells: different patterns of lymphokine secretion lead to different functional properties. Annual Review of Immunology, 7, p145-173*). Traditionally, TH1-type responses are associated with
5 the production of the INF- γ and IL-2 cytokines by T-lymphocytes. Other cytokines often directly associated with the induction of TH1-type immune responses are not produced by T-cells, such as IL-12. In contrast, TH2- type responses are associated with the secretion of IL-4, IL-5, IL-6 and IL-13.

10 It is known that certain vaccine adjuvants are particularly suited to the stimulation of either TH1 or TH2 - type cytokine responses. Traditionally the best indicators of the TH1:TH2 balance of the immune response after a vaccination or infection includes
direct measurement of the production of TH1 or TH2 cytokines by T lymphocytes *in vitro* after restimulation with antigen, and/or the measurement of the IgG1:IgG2a ratio
15 of antigen specific antibody responses.

Thus, a TH1-type adjuvant is one which preferentially stimulates isolated T-cell populations to produce high levels of TH1-type cytokines when re-stimulated with antigen *in vitro*, and promotes development of both CD8+ cytotoxic T lymphocytes and
20 antigen specific immunoglobulin responses associated with TH1-type isotype.

Adjuvants which are capable of preferential stimulation of the TH1 cell response are described in International Patent Application No. WO 94/00153 and WO 95/17209.

25 3 De-O-acylated monophosphoryl lipid A (3D-MPL) is one such adjuvant. This is known from GB 2220211 (Ribi). Chemically it is a mixture of 3 De-O-acylated monophosphoryl lipid A with 4, 5 or 6 acylated chains and is manufactured by Ribi Immunochem, Montana. A preferred form of 3 De-O-acylated monophosphoryl lipid A is disclosed in European Patent 0 689 454 B1 (SmithKline Beecham Biologicals SA).

Preferably, the particles of 3D-MPL are small enough to be sterile filtered through a 0.22micron membrane (European Patent number 0 689 454).

3D-MPL will be present in the range of 10µg - 100µg preferably 25-50µg per dose
5 wherein the antigen will typically be present in a range 2-50µg per dose.

Another preferred adjuvant comprises QS21, an Hplc purified non-toxic fraction derived from the bark of Quillaja Saponaria Molina. Optionally this may be admixed with 3 De-O-acylated monophosphoryl lipid A (3D-MPL), optionally together with a carrier.

10

The method of production of QS21 is disclosed in US patent No. 5,057,540.

Non-reactogenic adjuvant formulations containing QS21 have been described previously (WO 96/33739). Such formulations comprising QS21 and cholesterol have
15 been shown to be successful TH1 stimulating adjuvants when formulated together with an antigen.

Further adjuvants which are preferential stimulators of TH1 cell response include immunomodulatory oligonucleotides, for example unmethylated CpG sequences as
20 disclosed in WO 96/02555.

Combinations of different TH1 stimulating adjuvants, such as those mentioned hereinabove, are also contemplated as providing an adjuvant which is a preferential stimulator of TH1 cell response. For example, QS21 can be formulated together with
25 3D-MPL. The ratio of QS21 : 3D-MPL will typically be in the order of 1 : 10 to 10 : 1; preferably 1:5 to 5 : 1 and often substantially 1 : 1. The preferred range for optimal synergy is 2.5 : 1 to 1 : 1 3D-MPL: QS21.

Preferably a carrier is also present in the vaccine composition according to the invention. The carrier may be an oil in water emulsion, or an aluminium salt, such as aluminium phosphate or aluminium hydroxide.

- 5 A preferred oil-in-water emulsion comprises a metabolisable oil, such as squalene, alpha tocopherol and Tween 80. In a particularly preferred aspect the antigens in the vaccine composition according to the invention are combined with QS21 and 3D-MPL in such an emulsion. Additionally the oil in water emulsion may contain span 85 and/or lecithin and/or tricaprylin.

10

Typically for human administration QS21 and 3D-MPL will be present in a vaccine in the range of 1 µg - 200 µg, such as 10-100 µg, preferably 10 µg - 50 µg per dose.

- Typically the oil in water will comprise from 2 to 10% squalene, from 2 to 10% alpha tocopherol and from 0.3 to 3% tween 80. Preferably the ratio of squalene: alpha
15 tocopherol is equal to or less than 1 as this provides a more stable emulsion. Span 85 may also be present at a level of 1%. In some cases it may be advantageous that the vaccines of the present invention will further contain a stabiliser.

- 20 Non-toxic oil in water emulsions preferably contain a non-toxic oil, e.g. squalane or squalene, an emulsifier, e.g. Tween 80, in an aqueous carrier. The aqueous carrier may be, for example, phosphate buffered saline.

A particularly potent adjuvant formulation involving QS21, 3D-MPL and tocopherol in an oil in water emulsion is described in WO 95/17210.

25

The present invention also provides a polyvalent vaccine composition comprising a vaccine formulation of the invention in combination with other antigens, in particular antigens useful for treating cancers, autoimmune diseases and related conditions. Such a

polyvalent vaccine composition may include a TH-1 inducing adjuvant as hereinbefore described.

While the invention has been described with reference to certain BASB013 polypeptides and polynucleotides, it is to be understood that this covers fragments of the naturally occurring polypeptides and polynucleotides, and similar polypeptides and polynucleotides with additions, deletions or substitutions which do not substantially affect the immunogenic properties of the recombinant polypeptides or polynucleotides.

- 10 The antigen can also be delivered in the form of whole bacteria (dead or alive) or as subcellular fractions, these possibilities do include *N.meningitidis* itself.

Compositions, kits and administration

- 15 In a further aspect of the invention there are provided compositions comprising a BASB013 polynucleotide and/or a BASB013 polypeptide for administration to a cell or to a multicellular organism.

- The invention also relates to compositions comprising a polynucleotide and/or a polypeptide discussed herein or their agonists or antagonists. The polypeptides and polynucleotides of the invention may be employed in combination with a non-sterile or sterile carrier or carriers for use with cells, tissues or organisms, such as a pharmaceutical carrier suitable for administration to an individual. Such compositions comprise, for instance, a media additive or a therapeutically effective amount of a polypeptide and/or polynucleotide of the invention and a pharmaceutically acceptable carrier or excipient. Such carriers may include, but are not limited to, saline, buffered saline, dextrose, water, glycerol, ethanol and combinations thereof. The formulation should suit the mode of administration. The invention further relates to diagnostic and pharmaceutical packs and kits comprising one or more containers filled with one or more of the ingredients of the aforementioned compositions of the invention.

Polypeptides, polynucleotides and other compounds of the invention may be employed alone or in conjunction with other compounds, such as therapeutic compounds.

- 5 The pharmaceutical compositions may be administered in any effective, convenient manner including, for instance, administration by topical, oral, anal, vaginal, intravenous, intraperitoneal, intramuscular, subcutaneous, intranasal or intradermal routes among others.

- 10 In therapy or as a prophylactic, the active agent may be administered to an individual as an injectable composition, for example as a sterile aqueous dispersion, preferably isotonic.

- In a further aspect, the present invention provides for pharmaceutical compositions comprising a therapeutically effective amount of a polypeptide and/or polynucleotide, such as the soluble form of a polypeptide and/or polynucleotide of the present invention, agonist or antagonist peptide or small molecule compound, in combination with a pharmaceutically acceptable carrier or excipient. Such carriers include, but are not limited to, saline, buffered saline, dextrose, water, glycerol, ethanol, and combinations thereof. The invention further relates to pharmaceutical packs and kits comprising one or more containers filled with one or more of the ingredients of the aforementioned compositions of the invention.
- 15 20 Polypeptides, polynucleotides and other compounds of the present invention may be employed alone or in conjunction with other compounds, such as therapeutic compounds.

- 25 The composition will be adapted to the route of administration, for instance by a systemic or an oral route. Preferred forms of systemic administration include injection, typically by intravenous injection. Other injection routes, such as subcutaneous, intramuscular, or intraperitoneal, can be used. Alternative means for systemic administration include transmucosal and transdermal administration using penetrants such as bile salts or fusidic acids or other detergents. In addition, if a polypeptide or other compounds of the present

invention can be formulated in an enteric or an encapsulated formulation, oral administration may also be possible. Administration of these compounds may also be topical and/or localized, in the form of salves, pastes, gels, solutions, powders and the like.

- 5 For administration to mammals, and particularly humans, it is expected that the daily dosage level of the active agent will be from 0.01 mg/kg to 10 mg/kg, typically around 1 mg/kg. The physician in any event will determine the actual dosage which will be most suitable for an individual and will vary with the age, weight and response of the particular individual. The above dosages are exemplary of the average case. There can, of course,
10 be individual instances where higher or lower dosage ranges are merited, and such are within the scope of this invention.

The dosage range required depends on the choice of peptide, the route of administration, the nature of the formulation, the nature of the subject's condition, and the judgment of the
15 attending practitioner. Suitable dosages, however, are in the range of 0.1-100 μ g/kg of subject.

A vaccine composition is conveniently in injectable form. Conventional adjuvants may be employed to enhance the immune response. A suitable unit dose for vaccination is 0.5-5
20 microgram/kg of antigen, and such dose is preferably administered 1-3 times and with an interval of 1-3 weeks. With the indicated dose range, no adverse toxicological effects will be observed with the compounds of the invention which would preclude their administration to suitable individuals.

25 Wide variations in the needed dosage, however, are to be expected in view of the variety of compounds available and the differing efficiencies of various routes of administration. For example, oral administration would be expected to require higher dosages than administration by intravenous injection. Variations in these dosage levels can be adjusted

using standard empirical routines for optimization, as is well understood in the art.

Sequence Databases, Sequences in a Tangible Medium, and Algorithms

5 Polynucleotide and polypeptide sequences form a valuable information resource with which to determine their 2- and 3-dimensional structures as well as to identify further sequences of similar homology. These approaches are most easily facilitated by storing the sequence in a computer readable medium and then using the stored data in a known macromolecular structure program or to search a sequence database using well known searching tools, such
10 as the GCG program package.

Also provided by the invention are methods for the analysis of character sequences or strings, particularly genetic sequences or encoded protein sequences. Preferred methods of sequence analysis include, for example, methods of sequence homology analysis, such
15 as identity and similarity analysis, DNA, RNA and protein structure analysis, sequence assembly, cladistic analysis, sequence motif analysis, open reading frame determination, nucleic acid base calling, codon usage analysis, nucleic acid base trimming, and sequencing chromatogram peak analysis.

20 A computer based method is provided for performing homology identification. This method comprises the steps of: providing a first polynucleotide sequence comprising the sequence of a polynucleotide of the invention in a computer readable medium; and comparing said first polynucleotide sequence to at least one second polynucleotide or polypeptide sequence to identify homology.

25 A computer based method is also provided for performing homology identification, said method comprising the steps of: providing a first polypeptide sequence comprising the sequence of a polypeptide of the invention in a computer readable medium; and

comparing said first polypeptide sequence to at least one second polynucleotide or polypeptide sequence to identify homology.

All publications and references, including but not limited to patents and patent applications, cited in this specification are herein incorporated by reference in their entirety as if each individual publication or reference were specifically and individually indicated to be incorporated by reference herein as being fully set forth. Any patent application to which this application claims priority is also incorporated by reference herein in its entirety in the manner described above for publications and references.

DEFINITIONS

"Identity," as known in the art, is a relationship between two or more polypeptide sequences or two or more polynucleotide sequences, as the case may be, as determined by comparing the sequences. In the art, "identity" also means the degree of sequence relatedness between polypeptide or polynucleotide sequences, as the case may be, as determined by the match between strings of such sequences. "Identity" can be readily calculated by known methods, including but not limited to those described in (*Computational Molecular Biology*, Lesk, A.M., ed., Oxford University Press, New York, 1988; *Biocomputing: Informatics and Genome Projects*, Smith, D.W., ed., Academic Press, New York, 1993; *Computer Analysis of Sequence Data*, Part I, Griffin, A.M., and Griffin, H.G., eds., Humana Press, New Jersey, 1994; *Sequence Analysis in Molecular Biology*, von Heine, G., Academic Press, 1987; and *Sequence Analysis Primer*, Gribskov, M. and Devereux, J., eds., M Stockton Press, New York, 1991; and Carillo, H., and Lipman, D., *SIAM J. Applied Math.*, 48: 1073 (1988). Methods to determine identity are designed to give the largest match between the sequences tested. Moreover, methods to determine identity are codified in publicly available computer programs. Computer program methods to determine identity between two sequences include, but are not limited to, the GAP program in the GCG program package (Devereux, J., et al., *Nucleic Acids Research* 12(1):

387 (1984)), BLASTP, BLASTN (Altschul, S.F. et al., *J. Molec. Biol.* 215: 403-410 (1990), and FASTA(Pearson and Lipman Proc. Natl. Acad. Sci. USA 85; 2444-2448 (1988). The BLAST family of programs is publicly available from NCBI and other sources (*BLAST Manual*, Altschul, S., et al., NCBI NLM NIH Bethesda, MD 20894; 5 Altschul, S., et al., *J. Mol. Biol.* 215: 403-410 (1990). The well known Smith Waterman algorithm may also be used to determine identity.

Parameters for polypeptide sequence comparison include the following:

Algorithm: Needleman and Wunsch, *J. Mol Biol.* 48: 443-453 (1970)

10 Comparison matrix: BLOSSUM62 from Henikoff and Henikoff,
Proc. Natl. Acad. Sci. USA. 89:10915-10919 (1992)

Gap Penalty: 8

Gap Length Penalty: 2

A program useful with these parameters is publicly available as the "gap" program from 15 Genetics Computer Group, Madison WI. The aforementioned parameters are the default parameters for peptide comparisons (along with no penalty for end gaps).

Parameters for polynucleotide comparison include the following:

Algorithm: Needleman and Wunsch, *J. Mol Biol.* 48: 443-453 (1970)

20 Comparison matrix: matches = +10, mismatch = 0

Gap Penalty: 50

Gap Length Penalty: 3

Available as: The "gap" program from Genetics Computer Group, Madison WI. These 25 are the default parameters for nucleic acid comparisons.

A preferred meaning for "identity" for polynucleotides and polypeptides, as the case may be, are provided in (1) and (2) below.

(1) Polynucleotide embodiments further include an isolated polynucleotide comprising a polynucleotide sequence having at least a 50, 60, 70, 80, 85, 90, 95, 97 or 100% identity to the reference sequence of SEQ ID NO:1, wherein said polynucleotide sequence may be identical to the reference sequence of SEQ ID NO:1 or may include up to a certain integer number of nucleotide alterations as compared to the reference sequence, wherein said alterations are selected from the group consisting of at least one nucleotide deletion, substitution, including transition and transversion, or insertion, and wherein said alterations may occur at the 5' or 3' terminal positions of the reference nucleotide sequence or anywhere between those terminal positions, interspersed either individually among the nucleotides in the reference sequence or in one or more contiguous groups within the reference sequence, and wherein said number of nucleotide alterations is determined by multiplying the total number of nucleotides in SEQ ID NO:1 by the integer defining the percent identity divided by 100 and then subtracting that product from said total number of nucleotides in SEQ ID NO:1, or:

$$n_n \leq x_n - (x_n \bullet y),$$

wherein n_n is the number of nucleotide alterations, x_n is the total number of nucleotides in SEQ ID NO:1, y is 0.50 for 50%, 0.60 for 60%, 0.70 for 70%, 0.80 for 80%, 0.85 for 85%, 0.90 for 90%, 0.95 for 95%, 0.97 for 97% or 1.00 for 100%, and \bullet is the symbol for the multiplication operator, and wherein any non-integer product of x_n and y is rounded down to the nearest integer prior to subtracting it from x_n . Alterations of a polynucleotide sequence encoding the polypeptide of SEQ ID NO:2 may create nonsense, missense or frameshift mutations in this coding sequence and thereby alter the polypeptide encoded by the polynucleotide following such alterations.

By way of example, a polynucleotide sequence of the present invention may be identical to the reference sequence of SEQ ID NO:1, that is it may be 100% identical, or it may include up to a certain integer number of nucleic acid alterations as compared to the

reference sequence such that the percent identity is less than 100% identity. Such alterations are selected from the group consisting of at least one nucleic acid deletion, substitution, including transition and transversion, or insertion, and wherein said alterations may occur at the 5' or 3' terminal positions of the reference polynucleotide sequence or anywhere between those terminal positions, interspersed either individually among the nucleic acids in the reference sequence or in one or more contiguous groups within the reference sequence. The number of nucleic acid alterations for a given percent identity is determined by multiplying the total number of nucleic acids in SEQ ID NO:1 by the integer defining the percent identity divided by 100 and then subtracting that product from said total number of nucleic acids in SEQ ID NO:1, or:

$$n_n \leq x_n - (x_n \bullet y),$$

wherein n_n is the number of nucleic acid alterations, x_n is the total number of nucleic acids in SEQ ID NO:1, y is, for instance 0.70 for 70%, 0.80 for 80%, 0.85 for 85% etc., \bullet is the symbol for the multiplication operator, and wherein any non-integer product of x_n and y is rounded down to the nearest integer prior to subtracting it from x_n .

(2) Polypeptide embodiments further include an isolated polypeptide comprising a polypeptide having at least a 50, 60, 70, 80, 85, 90, 95, 97 or 100% identity to a polypeptide reference sequence of SEQ ID NO:2, wherein said polypeptide sequence may be identical to the reference sequence of SEQ ID NO:2 or may include up to a certain integer number of amino acid alterations as compared to the reference sequence, wherein said alterations are selected from the group consisting of at least one amino acid deletion, substitution, including conservative and non-conservative substitution, or insertion, and wherein said alterations may occur at the amino- or carboxy-terminal positions of the reference polypeptide sequence or anywhere between those terminal positions, interspersed either individually among the amino acids in the reference sequence or in one or more contiguous groups within the reference sequence, and wherein said number of

amino acid alterations is determined by multiplying the total number of amino acids in SEQ ID NO:2 by the integer defining the percent identity divided by 100 and then subtracting that product from said total number of amino acids in SEQ ID NO:2, or:

5 $n_a \leq x_a - (x_a \bullet y),$

wherein n_a is the number of amino acid alterations, x_a is the total number of amino acids in SEQ ID NO:2, y is 0.50 for 50%, 0.60 for 60%, 0.70 for 70%, 0.80 for 80%, 0.85 for 85%, 0.90 for 90%, 0.95 for 95%, 0.97 for 97% or 1.00 for 100%, and \bullet is the symbol for
10 the multiplication operator, and wherein any non-integer product of x_a and y is rounded down to the nearest integer prior to subtracting it from x_a .

By way of example, a polypeptide sequence of the present invention may be identical to the reference sequence of SEQ ID NO:2, that is it may be 100% identical, or it may
15 include up to a certain integer number of amino acid alterations as compared to the reference sequence such that the percent identity is less than 100% identity. Such alterations are selected from the group consisting of at least one amino acid deletion, substitution, including conservative and non-conservative substitution, or insertion, and wherein said alterations may occur at the amino- or carboxy-terminal positions of the
20 reference polypeptide sequence or anywhere between those terminal positions, interspersed either individually among the amino acids in the reference sequence or in one or more contiguous groups within the reference sequence. The number of amino acid alterations for a given % identity is determined by multiplying the total number of amino acids in SEQ ID NO:2 by the integer defining the percent identity divided by 100 and
25 then subtracting that product from said total number of amino acids in SEQ ID NO:2, or:

$$n_a \leq x_a - (x_a \bullet y),$$

wherein n_a is the number of amino acid alterations, x_a is the total number of amino acids in SEQ ID NO:2, y is, for instance 0.70 for 70%, 0.80 for 80%, 0.85 for 85% etc., and \bullet is the symbol for the multiplication operator, and wherein any non-integer product of x_a and y is rounded down to the nearest integer prior to subtracting it from x_a .

"Individual(s)," when used herein with reference to an organism, means a multicellular eukaryote, including, but not limited to a metazoan, a mammal, an ovid, a bovid, a simian, a primate, and a human.

"Isolated" means altered "by the hand of man" from its natural state, *i.e.*, if it occurs in nature, it has been changed or removed from its original environment, or both. For example, a polynucleotide or a polypeptide naturally present in a living organism is not "isolated," but the same polynucleotide or polypeptide separated from the coexisting materials of its natural state is "isolated", as the term is employed herein. Moreover, a polynucleotide or polypeptide that is introduced into an organism by transformation, genetic manipulation or by any other recombinant method is "isolated" even if it is still present in said organism, which organism may be living or non-living.

"Polynucleotide(s)" generally refers to any polyribonucleotide or polydeoxyribonucleotide, which may be unmodified RNA or DNA or modified RNA or DNA including single and double-stranded regions.

"Variant" refers to a polynucleotide or polypeptide that differs from a reference polynucleotide or polypeptide, but retains essential properties. A typical variant of a polynucleotide differs in nucleotide sequence from another, reference polynucleotide. Changes in the nucleotide sequence of the variant may or may not alter the amino acid sequence of a polypeptide encoded by the reference polynucleotide. Nucleotide changes may result in amino acid substitutions, additions, deletions, fusions and truncations in the polypeptide encoded by the reference sequence, as discussed below. A typical

variant of a polypeptide differs in amino acid sequence from another, reference polypeptide. Generally, differences are limited so that the sequences of the reference polypeptide and the variant are closely similar overall and, in many regions, identical. A variant and reference polypeptide may differ in amino acid sequence by one or more
5 substitutions, additions, deletions in any combination. A substituted or inserted amino acid residue may or may not be one encoded by the genetic code. A variant of a polynucleotide or polypeptide may be a naturally occurring such as an allelic variant, or it may be a variant that is not known to occur naturally. Non-naturally occurring variants of polynucleotides and polypeptides may be made by mutagenesis techniques
10 or by direct synthesis.

"Disease(s)" means any disease caused by or related to infection by a bacteria, including , for example, upper respiratory tract infection, invasive bacterial diseases, such as bacteremia and meningitis.

EXAMPLES:

The examples below are carried out using standard techniques, which are well known and routine to those of skill in the art, except where otherwise described in detail. The examples
5 are illustrative, but do not limit the invention.

**Example1: Discovery and confirmatory DNA sequencing of the BASB013 gene
from two *N.meningitidis* isolates.**

10 **A: BASB013 in *N. meningitidis* serogroup B strain ATCC13090.**

The BASB013 gene disclosed in SEQ ID NO:1 was first discovered in the Incyte
PathoSeq database containing unfinished genomic DNA sequences of the *N.
meningitidis* strain ATCC13090. The translation of the BASB013 polynucleotide
sequence, shown in SEQ ID NO:2, showed significant similarity (51% identity in a 456
15 amino acids overlap) to the MucD protein of *Pseudomonas aeruginosa*, the latter being
a homolog of the HtrA serine protease found in many bacteria, in particular in
Escherichia coli and *Haemophilus influenzae*.

The sequence of the BASB013 gene was further confirmed experimentally. For this
20 purpose, genomic DNA was extracted from 10^{10} cells of the *N.meningitidis* cells (strain
ATCC 13090) using the QIAGEN genomic DNA extraction kit (Qiagen GmbH), and
1 µg of this material was submitted to Polymerase Chain Reaction DNA amplification
using primers Hin-full-N (5'- GGA ATT CCA TAT GTT CAA AAA ATA CCA ATA
CC-3') [SEQ ID NO:11] and Hin-full-X (5'-CGC CGC TCG AGT TGC AGG TTT
25 AAT GCG ATG-3') [SEQ ID NO:12]. This PCR product was gel-purified and
subjected to DNA sequencing using the Big Dye Cycle Sequencing kit (Perkin-Elmer)
and an ABI 373A/PRISM DNA sequencer. DNA sequencing was performed on both
strands with a redundancy of 2 and the full-length sequence was assembled using the
Seqman program from the DNASTAR Lasergene software package. The resulting DNA

sequence and deduced peptide sequence are shown as SEQ ID NO:3 and SEQ ID NO:4 respectively. Three nucleotides in SEQ ID NO:3 (at positions 615 to 617) were found different from their counterpart in SEQ ID NO:1, as shown in Figure 1.

It should be noted that the start codon of the BASB013 gene is GTG, which is not
5 unusual in bacterial genes. That GTG codon has been translated as Methionine.

B: BASB013 in *N. meningitidis* serogroup B strain H44/76.

The sequence of the BASB013 gene was also determined in an other *N. meningitidis* serogroup B strain, the strain H44/76. For this purpose, genomic DNA was extracted
10 from the *N. meningitidis* strain H44/76 using the experimental conditions presented in Example 1. This material (1µg) was then submitted to Polymerase Chain Reaction DNA amplification using primers Hin-full-N and Hin-full-X specific for the BASB013 gene. A 1518 bp DNA fragment was obtained, digested by the *NdeI/XhoI* restriction endonucleases and inserted into the corresponding sites of the pET-24b
15 cloning/expression vector (Novagen) using standard molecular biology techniques (Molecular Cloning, a Laboratory Manual, Second Edition, Eds: Sambrook, Fritsch & Maniatis, Cold Spring Harbor press 1989). Recombinant pET-24b/BASB013 was then submitted to DNA sequencing using the Big Dyes kit (Applied biosystems) and analyzed on a ABI 373/A DNA sequencer in the conditions described by the supplier.
20 As a result, the polynucleotide and deduced polypeptide sequences, referred to as SEQ ID NO:5 and SEQ ID NO:6 respectively, were obtained.

Using the PILEUP program from the GCG package, a multiple alignment of the nucleotide sequences of SEQ ID NO:1, 3 and 5 was performed, and is displayed in
25 Figure 1. A pairwise comparison of identities is summarized in Table 1, showing that the three BASB013 polynucleotide gene sequences are all similar at a identity level greater than or equal to 96.2 %. Using the same PILEUP program from the GCG package, a multiple alignment of the protein sequences of SEQ ID NO:2, 4 and 6 was performed, and is displayed in Figure 2. A pairwise comparison of identities is

summarized in Table 2, showing that the three BASB013 polypeptide sequences are all similar at a identity level equal to or greater than 95.8 %.

Taken together, these data indicate strong sequence conservation of the BASB013 gene among the two *N.meningitidis* serogroup B strains.

5

Example 2: Expression and purification of recombinant BASB013 protein in Escherichia coli.

10 As represented in Figure 1, nucleotide sequence comparison between two *N. meningitidis* strains indicated that the first 1110 nucleotides of BASB013 shared a high degree of sequence conservation. In contrast, the last 390 nucleotides of the BASB013 gene showed some sequence variability. Consequently, expression vectors allowing the production of the full-length, the conserved (BASB013-C, [SEQ ID NO:7] and [SEQ ID NO:8]) or the variable (BASB013-V, [SEQ ID NO:9] and [SEQ ID NO:10]) part were
15 constructed. The construction of the pET-24b cloning/expression vector containing the full-length BASB013 gene was described in Example 1B. The BASB013-C fragment was amplified by PCR using oligonucleotide primers HC1 (5'-GAT ATA CAT ATG TTC AAA AAA TAC CAA TAC CTC-3') [SEQ ID NO:13] and HC2 (5'-CTA GGG CTC GAG TCC CGG CGT AAT GGC GCC GAC-3') [SEQ ID NO:14]. The
20 BASB013-V fragment was amplified by PCR using oligonucleotide primers HV1 (5'-GAT ATA CAT ATG AAA GAA GTC AGC CTC GGC GTA-3') [SEQ ID NO:15] and HV2 (5'-CTA GGG CTC GAG TTG CAG GTT TAA TGC GAT GAA-3') [SEQ ID NO:16]. Both PCR amplicons were digested using *NdeI* and *XhoI* and
25 inserted in the corresponding sites of the pET24b cloning/expression vector using standard molecular biology methodology.

For these constructs, the BASB013, the BASB013-C and the BASB013-V genes were isolated from the strain H44/76. These genes are introduced in fusion with a stretch of 6

histidine residues, and are placed under the control of the strong bacteriophage T7 gene 10 promoter. For expression study, this vector was introduced into the *Escherichia coli* strain BL21 DE3 (Novagen), in which, the gene for the T7 polymerase is placed under the control of the isopropyl-beta-D thiogalactoside (IPTG)-regulatable *lac* promoter.

- 5 Liquid cultures (100 ml) of the BL21 DE-3 [pET-24b/BASB013], BL21 DE-3 [pET-24b/BASB013-C] and BL21 DE-3 [pET-24b/BASB013-V] *E. coli* recombinant strains were grown at 37°C under agitation until the optical density at 600nm (OD600) reached 0.6. At that time-point, IPTG was added at a final concentration of 1mM and the culture was grown for 4 additional hours. The culture was then centrifuged at 10.000 rpm and
- 10 the pellet was frozen at -20°C for at least 10 hours. Subcellular localization of the polypeptide expressed from pET24b indicated that the BASB013 remained soluble in the bacterial cytoplasm whereas BASB013-C and BASB013-V were insoluble.
- Consequently, BASB013 was purified under mild, non-denaturing conditions whereas BASB013-C and BASB013-V were purified under denaturing conditions.

15

Purification of BASB013 under mild, non-denaturing conditions:

- After thawing, the cell paste was resuspended in 12.5 ml of start buffer (10mM sodium-phosphate pH7.4, NaCl 0.5M, Imidazole 10mM) containing 100mM of Pefablock (Boehringer Mannheim) protease inhibitor. The sample was then loaded at a flow-rate
- 20 of 1ml/min on a Ni²⁺ -loaded Hitrap column (Pharmacia Biotech). After passage of the flow trough, the column was washed succesively with start buffer (40ml) and start buffer containing 60mM Imidazole (30ml). The recombinant protein BASB013/His6 was then eluted from the column with 30ml of start buffer containing 500mM of imidazole and 3ml-size fractions were collected.

25

Purification of BASB013-C and BASB013-V under denaturing conditions:

After thawing, the cell paste was resuspended during 30 min at 25°C in buffer A (6M guanidine hydrochloride, 0.1M NaH₂PO₄, 0.01M Tris, pH 8.0), passed three-times through a needle and clarified by centrifugation (20000rpm, 15 min). The sample was

then loaded at a flow-rate of 1ml/min on a Ni²⁺-loaded Hitrap column (Pharmacia Biotech). After passage of the flowthrough, the column was washed successively with 40ml of buffer B (8M Urea, 0.1M NaH₂PO₄, 0.01M Tris, pH 8.0), 40ml of buffer C (8M Urea, 0.1M NaH₂PO₄, 0.01M Tris, pH 6.3). The recombinant protein

5 BASB013/His6 was then eluted from the column with 30ml of buffer C (8M Urea, 0.1M NaH₂PO₄, 0.01M Tris, pH 6.3) containing 500mM of imidazole and 3ml-size fractions were collected. As shown in Figure 3, enriched fractions (purity estimated to more than 80% pure in coomassie staining) were obtained for BASB013 (MW estimated to 52 kDa), and BSAB013-V (MW estimated to 14 kDa) after elution from

10 the column. For BASB013-C (MW estimated to 40 kDa), the recovery yield was very low and the protein was not detected by coomassie staining and required western blotting analysis (data not shown). These 3 polypeptides were reactive against a mouse monoclonal antibody raised against the 5-histidine motif. Taken together, these data indicate that the BASB013 gene can be expressed and purified under several

15 recombinant forms (BASB013/His6, BASB013-C/His6 and BASB013-V/His6) in *E.coli*.

Table 1. Pairwise identities of the BASB013 polynucleotide sequences (in %)

	SeqID No:3	SeqID No:5
SeqID No:1	99.8	96.2
SeqID No:3		96.4

Table 2. Pairwise identities of the BASB013 polypeptide sequences (in %)

	SeqID No:4	SeqID No:6
SeqID No:2	99.8	95.8
SeqID No:4		96.0

5

Example 3 : Presence of anti-BASB013 antibodies in sera from convalescent patients.

10

In this test, human convalescent sera have been tested by western-blotting for recognition of the purified recombinant BASB013 protein.

15 15µg of purified BASB013 protein (full length, cl. 8) are put into a SDS-PAGE gradient gel (4-20%, Novex, code n°EC60252) for electrophoretic migration. Proteins are transferred to nitrocellulose sheet (0.45µm, Bio-rad code n° 162-0114) at 100 volts for 1 hour using a Bio-rad Trans-blot system (code n°170-3930). The filter is then blocked with PBS - 0.05% Tween 20 overnight at room temperature, before incubation with the

20 human sera. These sera are diluted 100 fold in PBS - 0.05% Tween 20, and incubated on the nitrocellulose sheet for two hours at room temperature with gentle shaking, using a mini-blotter system (Miniprotean, Bio-rad code n° 170-4017). After three repeated washing steps in PBS - 0.05% Tween 20 for 5 min., the nitrocellulose sheet is incubated at room temperature for 1 hour under gentle shaking with the appropriate conjugate

25 (biotinylated anti-human Ig antibodies from sheep, Amersham code n°RPN1003) diluted at 1/500 in the same washing buffer. The membrane is washed three times as

previously, and incubated for 30 min. with agitation using the streptavidin-peroxidase complex (Amersham code n°1051) diluted at 1/1000 in the washing buffer. After the last three repeated washing steps, the revelation occurs during the 20 min incubation time in a 50ml solution containing 30mg 4-chloro-1-naphtol (Sigma), 10ml methanol, 5 40ml of ultra-pure water, and 30µl of H₂O₂. The staining is stopped while washing the membrane several times in distilled water. In part A of the western-blot, revelation was done with a pool of mice sera from animals immunized with 2µg of Outer Membrane Proteins (OMP, obtained by Lithium Chloride extraction) injected in SB62 with 5µg MPL and 2µg QS21 by the subcutaneous and intraperitoneal routes on days 0, 10 21 and 42. Animals were bled on day 49. Mice antibodies were detected as for human antibodies except the conjugate used was a biotinylated anti-mouse Ig antibodies from sheep, Amersham code n°RPN1001) diluted at 1/500.

Results illustrated in Figure 4 (Part B) show that 6/7 convalescent sera tested react against the purified recombinant BASB013 protein (see arrow; only patient n° 261324 15 does not show any reactivity against the protein). The BASB013 band is clearly visible at around 53 kD. In part A of the western-blot, it is seen that mice immunized with OMP from strain H44/76 recognize very well the recombinant BASB013 polypeptide at the same molecular weight (53 kD).

Deposited materials

A deposit containing a *Neisseria meningitidis* Serogroup B strain has been deposited with the American Type Culture Collection (herein "ATCC") on June 22, 1997 and assigned deposit number 13090. The deposit was described as *Neisseria Meningitidis* (Albrecht and Ghon) and is a freeze-dried, 1.5-2.9 kb insert library constructed from *N. meningitidis* isolate. The deposit is described in Int. Bull. Bacteriol. Nomencl. Taxon. 8: 1-15 (1958).

The *Neisseria meningitidis* strain deposit is referred to herein as "the deposited strain" or as "the DNA of the deposited strain."

The deposited strain contains the full length BASB013 gene. The sequence of the polynucleotides contained in the deposited strain, as well as the amino acid sequence of any polypeptide encoded thereby, are controlling in the event of any conflict with any description of sequences herein.

The deposit of the deposited strain has been made under the terms of the Budapest Treaty on the International Recognition of the Deposit of Micro-organisms for Purposes of Patent Procedure. The strain will be irrevocably and without restriction or condition released to the public upon the issuance of a patent. The deposited strain is provided merely as convenience to those of skill in the art and is not an admission that a deposit is required for enablement, such as that required under 35 U.S.C. §112.

INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No

PCT/EP 99/02765

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
EP 0301992 A	01-02-1989	AT 122893 T	15-06-1995
		AU 615461 B	03-10-1991
		AU 2031288 A	25-05-1989
		AU 5319794 A	24-03-1994
		AU 706213 B	10-06-1999
		AU 7422696 A	20-02-1997
		AU 8134991 A	31-10-1991
		DE 3853854 D	29-06-1995
		DE 3853854 T	08-02-1996
		ES 2074445 T	16-09-1995
		GR 3017218 T	30-11-1995
		IN 167607 A	24-11-1990
		JP 1125328 A	17-05-1989
		RU 2023448 C	30-11-1994
		US 5597572 A	28-01-1997
		US 5747653 A	05-05-1998
WO 9802547 A	22-01-1998	FR 2751000 A	16-01-1998
		AU 3697797 A	09-02-1998

INTERNATIONAL SEARCH REPORT

Internal Application No

PCT/EP 99/02765

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	<p>WO 98 02547 A (INST NAT SANTE RECH MED ;MAX PLANCK GESELLSCHAFT (DE); SMITHKLINE) 22 January 1998 (1998-01-22) page 22, line 9 - line 26; claims 1-32; examples 1,5,6</p> <p>-----</p>	1-28

INTERNATIONAL SEARCH REPORT

International Application No

PCT/EP 99/02765

A. CLASSIFICATION OF SUBJECT MATTER

IPC 6 C12N15/31 C12N15/11 C12N15/70 C12N1/21 C07K16/12
G01N33/53 G01N33/50 A61K31/70 A61K39/095 A61K39/40
C07K14/22

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC 6 C07K C12N

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	LOOSMORE ET AL.: "The Haemophilus influenza HtrA protein is a protective antigen"	6,21, 24-26,28
A	INFECTION AND IMMUNITY, vol. 66, no. 3, March 1998 (1998-03), pages 899-906, XP002112241 the whole document	1-5, 7-20, 22-24
A	EP 0 301 992 A (NACIONAL DE BIOPREPARADOS CENT) 1 February 1989 (1989-02-01) page 3, line 43 - line 62 page 8, line 36 - page 10, line 44 --- -/--	1-28



Further documents are listed in the continuation of box C.



Patent family members are listed in annex.

* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search

17 August 1999

Date of mailing of the international search report

30/08/1999

Name and mailing address of the ISA

European Patent Office, P.B. 5818 Patentlaan 2
NL - 2280 HV Rijswijk
Tel. (+31-70) 340-2040, Tx. 31 651 epo nl.
Fax: (+31-70) 340-3016

Authorized officer

van Klompenburg, W

KEVSLGVWRKGEEITIKVKLGNAAEHIGASSKTDEAPYTEQQSGTFSVESAGITLQTHT
DSSGGHLVVVRVSDAAERAGLRGDEILAVGQVPVNDEAGFRKAMDKAGKNVPLLIMRR
GNTLFIALNLQ

SEQ ID NO:11

GGA ATT CCA TAT GTT CAA AAA ATA CCA ATA CC

SEQ ID NO:12

CGC CGC TCG AGT TGC AGG TTT AAT GCG ATG

SEQ ID NO:13

GAT ATA CAT ATG TTC AAA AAA TAC CAA TAC CTC

SEQ ID NO:14

CTA GGG CTC GAG TCC CGG CGT AAT GGC GCC GAC

SEQ ID NO:15

GAT ATA CAT ATG AAA GAA GTC AGC CTC GGC GTA

SEQ ID NO:16

CTA GGG CTC GAG TTG CAG GTT TAA TGC GAT GAA

VPLLIMRRGNTLFIALNLQ

SEQIDN° 7

ATGTTCAAAAAATACCAATACCTCGCTTTGGCAGCACTGTGTGCCGCTCGCTGGCAGGCTGCGACAAGGCAGGCAGC
TTTTTCGGTGCGGACAAAAAAGAAGCATCCTTCGTAGAACGCATCGAACACACCAAAGACGCGCAGCGTCAGTATG
CTGCTGCCCCGACTTTGcCCAACCTGGTTCAAAGTGAAGGTCCGGCAGTCGTCAATATTCAGGCAGCCCCCGCCCCGCGC
ACCCAAAACGGCAGCGGCAATGCCGAAAACGATTCCGACCCGATTGCCGACAACGACCCGTTCTACGAATTTTCAAA
CGCCTCGTCCCGAATATGCCCGAAATCCCCAAGAAGAAGCAGATGACGGCGGATTGAACTTCGGTTTCGGGCTTCATC
ATCAGCAAAGACGGCTACATCCTGACCAATACCCACGTCGTTACCGGCATGGGCAGTATCAAAGTCCTGCTCAACGAC
AAGCGGAATATACCGCCAAACTCATCGTTTCGGATGTCCAATCCGATGTGCGCCTTCTGAAAATCGACGCAACGGAA
GAGCTGCCCCGTCGTCAAATCGGCAATCCCAAAGATTTGAAACCGGGCGAATGGGTGCGCGCCATCGGCGCGCCCTTC
GGCTTCGACAACAGCGTGACCGCGGCATCGTGTCCGCCAAAGGCAGAAGCCTGCCCAACGAAAGCTACACACCCTTC
ATCCAAACCGACGTTGCCATCAATCCGGGCAACTCCGGCGGCCCGCTGTTCAACTTAAAGGACAGGTCGTCCGCATC
AACTCGCAAATATACAGCCGACGCGCGGATTTCATGGGCATTTCTTCGCCATCCCGATTGACGTTGCCATGAATGTC
GCCGAACAGCTGAAAAACACCGGCAAAGTCCAACGCGGACAACCTGGGCGTGATTATTCAAGAAGTATCCTACGGTTTG
GCACAATCGTTCCGtTTGGACAAAGCCGGCGGCGCACTGATTGCCAAAATCCTGCCCGGCAGCCCCGAGAACGTGCC
GGCCTGCAGGCGGGCGACATCGTCCTCAGCCTCGACGGCGGAGAAATACGTTCTTCCGGCGACCTTCCCCTTATGGTC
GGCGCCATTACGCCGGA

SEQIDN° 8

MFKKYQYLALALCAASLAGCDKAGSFFGADKKEASFVERIEHTKDDGSVSMLLPDLFAQ
LVQSEGPVAVNIQAAPAPRTQNGSGNAENDSDPIADNDPFYEFFKRLVPMPEI PQEEA
DDGGLNFGSGFIISKDGYILTNTHVVTGMGSIKVLLNDKREYTAKLIGSDVQSDVALLK
IDATEELPVVKIGNPKDLKPGEWVAAIGAPFGFDNSVTAGIVSAKGRSLPNESYTPFIQ
TDVAINPGNSGGPLFNLKGQVVGINSQIYSRSGGFMGISFAIPIDVAMNVAEQLKNTGK
VQRGQLGVI IQEVSYGLAQSFGLDKAGGALIAKILPGSPAERAGLQAGDIVLSLDGGEI
RSSGDLPMVMGAITPG

SEQIDN° 9

AAAGAAGTCAGCCTCGGCGTATGGCGCAAAGGCGAAGAAATCACAATCAAAGTCAAGCT
GGGCAAcGCCGCCGAGCATATCGGCGCATCATCCAAAACAGATGAAGCCCCCTACACCG
AACAGCAATCCGGTACGTTCTCGGTGCAATCCGCAGGCATTACCCTTCAGACACATACC
GACAGCAGCGGCGGACACCTCGTCGTGCTACGGGTTTCCGACGCGGCAGAACGCGCAGG
CTTGAGGCGCGGCGACGAAATTCCTTGCCGTCGGGCAAGTCCCCGTCAATGACGAAGCCG
GTTTCCGCAAAGCTATGGACAAGGCAGGCAAAAACGTCCTCCCTGCTGATCATGCGCCGT
GGCAACACGCTGTTTCATCGCATTAACCTGCAATAA

SEQIDN° 10

INPGNSGGPLFNLKGQVVGINSQIYSRSGGMGISFAIPIDVAMNVAEQLKNTGKVQRGQ
 LGVIIQEVSYGLAQSFGLDKAGGALIAKILPGSPAERAGLQAGDIVLSLDGGEIRSSGDL
 PVMVGAITPGKEVSLGVWRKGKEITVAVKLGNASEQTGSSEPDKAPYAEHQSGTFSVES
 AGITLQTHTDSSGGRLVVVRVSGAAERAGLRRGDEILAVGQVPVNDEDGFRKAMDKAGKN
 VPLLVMRRGNTLFIALNLQ

SEQ ID NO:5

Neisseria meningitidis BASB013 polynucleotide sequence from strain H44/76

GTGTTCAAAAAATACCAATACCTCGCTTTGGCAGCACTGTGTGCCGCTCGCTGGCAGGCTGCGACAAGGCAGGCAGC
 TTTTTCGGTGCGGACAAAAAGAAGCATCCTTCGTAGAACGCATCGAACACACCAAAGACGACGGCAGCGTCAGTATG
 CTGCTGCCCCGACTTTGcCCAACCTGGTTCAAAGTGAAGGTCCGGCAGTCGTCAATATTCAGGCAGCCCCGCCCGCGC
 ACCCAAAACGCGCAGCGGCAATGCCGAAAACGATTCCGACCCGATTGCCGACAACGACCCGTTCTACGAATTTTTCAA
 CGCCTCGTCCCGAATATGCCCGAAAATCCCCAAGAAGAAGCAGATGACGGCGGATTGAACTTCGGTTCGGGCTTCATC
 ATCAGCAAAGACGGCTACATCCTGACCAATACCCACGTCGTTACCGGCATGGGCAGTATCAAAGTCTGCTCAACGAC
 AAGCGCGAATATACCGCCAAACTCATCGGTTCCGATGTCCAATCCGATGTGCCCTTCTGAAAATCGACGCAACGGAA
 GAGCTGCCCCGTGCTCAAAATCGGCAATCCCAAAGATTGAAACCGGGCGAATGGGTGCGCGCCATCGGCGCGCCCTTC
 GGCTTCGACAACAGCGTGACCGCCGGCATCGTGTCCGCCAAAGGCAGAAGCCTGCCcAÀCGAAAGCTACACACCCTTC
 ATCCAAACCGACGTTGCCATCAATCCGGGCAACTCCGGCGGCCCGCTGTTCAACTTAAAGGACAGGTCGTCCGCATC
 AACTCGCAAATATACAGCCGAGCGGCGGATTATGGGCATTTCTTCGCCATCCCGATTGACGTTGCCATGAATGTC
 GCCGAACAGCTGAAAAACACCGGCAAAGTCCAACGCGGACAACCTGGGCGTGATTATTCAAGAAGTATCCTACGTTTG
 GCACAATCGTTCCGcTTGGACAAAGCCGGCGGCGCACTGATTGCCAAAATCCTGCCCGGACCCCCGAGAACGTGCC
 GGCTGCGAGGCGGGGACATCGTCCTCAGCCTCGACGGCGGAGAAATACGTTCTTCGGCGACCTTCCCCTTATGGTC
 GGCGCCATTACGCCGGGAAAAGAAGTCAGCCTCGCGGTATGGCGCAAAGGCGAAGAAATCACAATCAAAGTCAAGCTG
 GGCAAcGCCGCCGAGCATATCGGCGcATCATCCAAAACAGATGAAGCCCCCTACACCGAACAGCAATCCGGTACGTTT
 TCGGTGCAATCCGCAGGCATTACCTTTCAGACACATACCGACAGCAGCGCGGACACCTCGTCGTACGGGTTTCC
 GACGCGGCAGAACGCGCAGGCTTGAGGCGCGGCGACGAAATTCTTGCCGTGCGGCAAGTCCCCGTCAATGACGAAGCC
 GGTTTCCGCAAAGCTATGGACAAGGCAGGCAAAAACGTCCCCCTGCTGATCATGCGCCGTGGCAACACGCTGTTTCATC
 GCATTAAACCTGCAATAA

SEQ ID NO:6

Neisseria meningitidis BASB013 polypeptide sequence deduced from the
 polynucleotide sequence of SEQ ID NO:5

MFKKYQYLALALCAASLAGCDKAGSFFGADKKEASFVERIEHTKDDGSVSMMLLPDFAQL
 VQSEGPVVNIQAAPAPRTQNGSGNAENDSDPIADNDPFYEFFKRLVPMPEIPQEEADD
 GGLNFGSGFIISKDGYILTNTHVVTGMGSIKVLLNDKREYTAKLIGSDVQSDVALLKIDA
 TEELPVVKIGNPKDLKPGEWVAAIGAPFGFDNSVTAGIVSAKGRSLPNESYTPFIQTDVA
 INPGNSGGPLFNLKGQVVGINSQIYSRSGGMGISFAIPIDVAMNVAEQLKNTGKVQRGQ
 LGVIIQEVSYGLAQSFGLDKAGGALIAKILPGSPAERAGLQAGDIVLSLDGGEIRSSGDL
 PVMVGAITPGKEVSLGVWRKGEEITIKVKGNAAEHIGASSKTDEAPYTEQQSGTFSVES
 AGITLQTHTDSSGHLVVVRVSDAAERAGLRRGDEILAVGQVPVNDEAGFRKAMDKAGKN

INPGNSGGPLFNLKGQVVGINSQIYSRSGGFMGISFAIPIDVAMNVAEQLKNTGKVQRGQ
 LGVIIQEVSYGLAQSFGLDKAGGALIAKILPGSPAERAGLQAGDIVLSLDGGEIRSSGDL
 PVMVGAITPGKEVSLGVWRKGKEITVAVKLGNASEQTGSSSEPDKAPYAEHQSGTFSVES
 AGITLQTHTDSSGGRLVVVRVSGAAERAGLRRGDEILAVGQVPVNDEDGFRKAMDKAGKN
 VPLLVMRRGNTLFIALNLQ

SEQ ID NO:3

Neisseria meningitidis BASB013 polynucleotide sequence (sequenced from strain
 ATCC 13090)

GTGTTCAAAAAATACCAATACCTCGCTTTGGCAGCACTGTGTGCCGCCTCGCTGGCAGGC
 TGCGACAAAGCCGGCAGCTTTTCGGTGCGGACAAAAAGAAGCATCCTTCGTAGAACGC
 ATCGAACACACCAAAGACGACGGCAGCGTCAGTATGCTGCTGCCCCGACTTTGTCCAACGTG
 GTTCAAAGCGAAGGCCCGGAGTCGTCAATATTCAGGCAGCCCCCGCCCCGCGCACCCAA
 AACGGCAGCGGCAATGCCGAAACCGATTCCGACCCGCTTGCCGACAGCGACCCGTTCTAC
 GAATTTTCAAACGCCTCGTCCCGAATATGCCCGAAATCCCCAAGAAGAAGCAGATGAC
 GGCGGATTGAACTTEGGTTCTGGGCTTCATCATCAGCAAAGACGGCTATATTCTGACCAAT
 ACGCACGTCGTTACCGGCATGGGCAGTATCAAAGTCTTGCTCAACGACAAGCGCGAATAT
 ACCGCCAAACTCATCGGTTCCGGATGTCCAATCCGATGTCGCCCTTCTGAAAATCGACGCA
 ACGGAAGAGCTGCCCCGTCGTCAAAATCGGCAATCCCAAAGATTGAAACCGGGCGAATGG
 GTCGCCGCCATCGGCGCGCCCTTCGGCTTCGACAAACAGCGTGACCGCCGGCATCGTGTCC
 GCCAAAGGCAGAAGCCTGCCCAACGAAAGCTACACACCCTTCATCCAAACCGACGTTGCC
 ATCAATCCGGGCAACTCCGGCGGCCCGCTGTTCAACCTGAAAGGACAGGTGCTCGGCATC
 AACTCGCAAATATACAGCCGACGCGCGGATTTCATGGGCATTTCTTCGCCATCCCGATT
 GACGTTGCCATGAATGTGCGCGAACAGCTGAAAAACACCGGCAAAGTCCAACGCGGACAA
 CTGGGCGTGATTATTCAAGAAGTATCCTACGGTTTGGCACAATCGTTTCGGTTTGGACAAA
 GCCGGCGGCGCACTGATTGCCAAAATCCTGCCCGGACCCCCGAGAACGTGCCGGCCTG
 CAGGCGGGCGACATCGTCCTCAGCCTCGACGGCGGAGAAATACGTTCTTCGGGCGACCTT
 CCCGTTATGGTTCGGCGCCATTACGCCGGGAAAAGAAGTCAGCCTCGGCGTATGGCGCAAA
 GGTAAGGAAATCACGTTGCCGTCAAACCTGGGCAATGCTTCCGAACAAACCGGTTCTCG
 TCCGAGCCGGACAAAGCCCCCTTATGCCGAACACCAATCCGGTACGTTCTCGGTGCAATCC
 GCAGGCATTACCCTTCAGACACATACCGACAGCAGCGCGGACGGCTTGTGTCGTGCGG
 GTTTCGGGGGCGGCAGAACGCGCAGGCTTGAGGCGCGGCGACGAAATCCTTGCCGTCGGG
 CAAGTCCCCGTCAATGACGAAGACGGTTTCCGCAAAGCTATGGACAAGGCAGGCAAAAAC
 GTCCCCCTGCTGGTCATGCGCCGTGGCAACACGCTGTTTCATCGCATTAAACCTGCAATAA

SEQ ID NO:4

Neisseria meningitidis BASB013 polypeptide sequence deduced from the
 polynucleotide sequence of SEQ ID NO:3

MFKKYQYLALCAASLAGCDKAGSFFGADKKEASFVERIEHTKDDGSVSMLLPDFVQL
 VQSEGPVAVNIQAAPAPRTQNGSNAETDSDPLADSDPFYEFFKRLVPMPEIPQEEADD
 GGLNFGSGFIISKDGYILTNHVVTGMGSIKVLLNDKREYTAKLIGSDVQSDVALLKIDA
 TEELPVVKIGNPKDLKPGEWAAIGAPFGFDNSVTAGIVSAKGRSLPNESYTPFIQTDVA

BASB013 Polynucleotide and Polypeptide Sequences**SEQ ID NO:1**

Neisseria meningitidis BASB013 polynucleotide sequence (strain ATCC 13090)

GTGTTCAAAAAATACCAATACCTCGCTTTGGCAGCACTGTGTGCCGCTCGCTGGCAGGC
TGCGACAAAGCCGGCAGCTTTTTTCGGTGCGGACAAAAAGAAGCATCCTTCGTAGAACGC
ATCGAACACACCAAAGACGACGGCAGCGTCAGTATGCTGCTGCCCCGACTTTGTCCAAC TG
GTTCAAAGCGAAGGCCCGGCAGTCGTCAATATT CAGGCAGCCCCCGCCCCGCGCACCCAA
AACGGCAGCGGCAATGCCGAAACCGATTCCGACCCGCTTGCCGACAGCGACCCGTTCTAC
GAATTTTTCAAACGCCTCGTCCCGAATATGCCCCGAAATCCCCAAGAAGAAGCAGATGAC
GGCGGATTGAACTTCGGTTTCGGGCTTCATCATCAGCAAAGACGGCTATATTCTGACCAAT
ACGCACGTCGTTACCGGCATGGGCAGTATCAAAGTCTGCTCAACGACAAGCGCGAATAT
ACCGCCAAACTCATCGGTTTCGGATGTCCAATCCGATGTGCGCCTTCTGAAAATCGACGCA
ACGGAAGAGCTGCCCCGTCGTCAAATCGGCAATCCCAAAGATTGAAACCGGGCGAATGG
GTCGCGCCCATCGGGCGGCCCTTCGGCTTCGACAACAGCGTGACCGCCGGCATCGTGTCC
GCCAAAGGCAGAAGCCTGCCCAACGAAAGCTACACACCCTTCATCCAAACCGACGTTGCC
ATCAATCCGGGCAACTCCGGCGGCCCGCTGTTCAACCTGAAAGGACAGGTCGTGCGCATC
AACTCGCAAATATACAGCCGACGCGCGGATTTCATGGGCATTTCTTCGCCATCCCGATT
GACGTTGCCATGAATGTGCGCGAACAGCTGAAAAACACCGGCAAAGTCCAACGCGGACAA
CTGGGCGTGATTATTCAAGAAGTATCCTACGGTTTGGCACAATCGTTTCGGTTTGGACAAA
GCCGGCGGCGCACTGATTGCCAAAATCCTGCCCGGCAGCCCCGAGAACGTGCCGGCCTG
CAGGCGGGCGACATCGTCTCAGCCTCGACGGCGGAGAAATACGTTCTTCGGGCGACCTT
CCCGTTATGGTTCGGCGCCATTACGCCGGGAAAAGAAGTCAGCCTCGGCGTATGGCGCAA
GGTAAGGAAATCACCGTTGCCGTCAAACCTGGGCAATGCTTCGGAACAAACCGGTTCTCG
TCCGAGCCGGACAAAGCCCCCTTATGCCGAACACCAATCCGGTACGTTCTCGGTCGAATCC
GCAGGCATTACCCTTCAGACACATACCGACAGCAGCGGCGGACGGCTTGTGTCGTGCGG
GTTTCGGGGGCGGCAGAACGCGCAGGCTTGAGGCGCGGCGACGAAATCCTTGCCGTGCGG
CAAGTCCCCGTCAATGACGAAGACGGTTTCGCGAAAGCTATGGACAAGGCAGGCAAAAAC
GTCCCCCTGCTGGTCATGCGCCGTGGCAACACGCTGTTTCATCGCATTAAACCTGCAATAA

SEQ ID NO:2

Neisseria meningitidis BASB013 polypeptide sequence deduced from the
polynucleotide sequence of SEQ ID NO:1

MFKKYQYLALCAASLAGCDKAGSFFGADKKEASFVERIEHTKDDGSVSMLLPDFVQL
VQSEGPAVVNIQAAPAPRTQNGSGNAETDSDPLADSDPFYEFFKRLVPMPEIPQEEADD
GGLNFGSGFIISKDGYILTNHVVTMGMSIKVLLNDKREYTAKLIGSDVQSDVALLKIDA
TEELPVVKIGNPKDLKPGEWVAAIGRPFPGDNSVTAGIVSAKGRSLPNESYTPFIQTDVA

<223> primer

<400> 16

ctagggctcg agttgcaggt ttaatgcgat gaa

33

<210> 13

<211> 33

<212> DNA

<213> Artificial Sequence

<220>

<223> primer

<400> 13

gatatacata tgttcaaaaa ataccaatac ctc

33

<210> 14

<211> 33

<212> DNA

<213> Artificial Sequence

<220>

<223> primer

<400> 14

ctagggctcg agtcccggcg taatggcgcc gac

33

<210> 15

<211> 33

<212> DNA

<213> Artificial Sequence

<220>

<223> primer

<400> 15

gatatacata tgaaagaagt cagcctcggc gta

33

<210> 16

<211> 33

<212> DNA

<213> Artificial Sequence

<220>

Lys Val Lys Leu Gly Asn Ala Ala Glu His Ile Gly Ala Ser Ser Lys
 20 25 30
 Thr Asp Glu Ala Pro Tyr Thr Glu Gln Gln Ser Gly Thr Phe Ser Val
 35 40 45
 Glu Ser Ala Gly Ile Thr Leu Gln Thr His Thr Asp Ser Ser Gly Gly
 50 55 60
 His Leu Val Val Val Arg Val Ser Asp Ala Ala Glu Arg Ala Gly Leu
 65 70 75 80
 Arg Arg Gly Asp Glu Ile Leu Ala Val Gly Gln Val Pro Val Asn Asp
 85 90 95
 Glu Ala Gly Phe Arg Lys Ala Met Asp Lys Ala Gly Lys Asn Val Pro
 100 105 110
 Leu Leu Ile Met Arg Arg Gly Asn Thr Leu Phe Ile Ala Leu Asn Leu
 115 120 125
 Gln

<210> 11

<211> 32

<212> DNA

<213> Artificial Sequence

<220>

<223> Primer

<400> 11

ggaattccat atgttcaaaa aataccaata cc

32

<210> 12

<211> 30

<212> DNA

<213> Artificial Sequence

<220>

<223> primer

<400> 12

cgccgctcga gttgcaggtt taatgcgatg

30

Val Val Gly Ile Asn Ser Gln Ile Tyr Ser Arg Ser Gly Gly Phe Met
 260 265 270
 Gly Ile Ser Phe Ala Ile Pro Ile Asp Val Ala Met Asn Val Ala Glu
 275 280 285
 Gln Leu Lys Asn Thr Gly Lys Val Gln Arg Gly Gln Leu Gly Val Ile
 290 295 300
 Ile Gln Glu Val Ser Tyr Gly Leu Ala Gln Ser Phe Gly Leu Asp Lys
 305 310 315 320
 Ala Gly Gly Ala Leu Ile Ala Lys Ile Leu Pro Gly Ser Pro Ala Glu
 325 330 335
 Arg Ala Gly Leu Gln Ala Gly Asp Ile Val Leu Ser Leu Asp Gly Gly
 340 345 350
 Glu Ile Arg Ser Ser Gly Asp Leu Pro Val Met Val Gly Ala Ile Thr
 355 360 365
 Pro Gly
 370

<210> 9

<211> 390

<212> DNA

<213> Bacteria

<400> 9

```

aaagaagtca gcctcggcgt atggcgcaaa ggcgaaagaaa tcacaatcaa agtcaagctg      60
ggcaacgccg ccgagcatat cggcgcatca tccaaaacag atgaagcccc ctacaccgaa      120
cagcaatccg gtacgtttctc ggtcgaatcc gcaggcatta cccttcagac acataccgac      180
agcagcgggc gacacctcgt cgtcgtacgg gtttcggacg cggcagaacg cgcaggcttg      240
aggcgcgggc acgaaattct tgccgtcggg caagtccccg tcaatgacga agccgggtttc      300
cgcaaagcta tggacaaggc aggcaaaaac gtccccctgc tgatcatgcg ccgtggcaac      360
acgctgttca tcgcattaaa cctgcaataa                                     390
  
```

<210> 10

<211> 129

<212> PRT

<213> Bacteria

<400> 10

Lys Glu Val Ser Leu Gly Val Trp Arg Lys Gly Glu Glu Ile Thr Ile
 1 5 10 15

<210> 8

<211> 370

<212> PRT

<213> Bacteria

<400> 8

```

Met Phe Lys Lys Tyr Gln Tyr Leu Ala Leu Ala Ala Leu Cys Ala Ala
 1           5           10           15
Ser Leu Ala Gly Cys Asp Lys Ala Gly Ser Phe Phe Gly Ala Asp Lys
      20           25           30
Lys Glu Ala Ser Phe Val Glu Arg Ile Glu His Thr Lys Asp Asp Gly
      35           40           45
Ser Val Ser Met Leu Leu Pro Asp Phe Ala Gln Leu Val Gln Ser Glu
      50           55           60
Gly Pro Ala Val Val Asn Ile Gln Ala Ala Pro Ala Pro Arg Thr Gln
65           70           75           80
Asn Gly Ser Gly Asn Ala Glu Asn Asp Ser Asp Pro Ile Ala Asp Asn
      85           90           95
Asp Pro Phe Tyr Glu Phe Phe Lys Arg Leu Val Pro Asn Met Pro Glu
      100          105          110
Ile Pro Gln Glu Glu Ala Asp Asp Gly Gly Leu Asn Phe Gly Ser Gly
      115          120          125
Phe Ile Ile Ser Lys Asp Gly Tyr Ile Leu Thr Asn Thr His Val Val
      130          135          140
Thr Gly Met Gly Ser Ile Lys Val Leu Leu Asn Asp Lys Arg Glu Tyr
      145          150          155          160
Thr Ala Lys Leu Ile Gly Ser Asp Val Gln Ser Asp Val Ala Leu Leu
      165          170          175
Lys Ile Asp Ala Thr Glu Glu Leu Pro Val Val Lys Ile Gly Asn Pro
      180          185          190
Lys Asp Leu Lys Pro Gly Glu Trp Val Ala Ala Ile Gly Ala Pro Phe
      195          200          205
Gly Phe Asp Asn Ser Val Thr Ala Gly Ile Val Ser Ala Lys Gly Arg
      210          215          220
Ser Leu Pro Asn Glu Ser Tyr Thr Pro Phe Ile Gln Thr Asp Val Ala
      225          230          235          240
Ile Asn Pro Gly Asn Ser Gly Gly Pro Leu Phe Asn Leu Lys Gly Gln
      245          250          255

```

405 410 415
 Ser Val Glu Ser Ala Gly Ile Thr Leu Gln Thr His Thr Asp Ser Ser
 420 425 430
 Gly Gly His Leu Val Val Val Arg Val Ser Asp Ala Ala Glu Arg Ala
 435 440 445
 Gly Leu Arg Arg Gly Asp Glu Ile Leu Ala Val Gly Gln Val Pro Val
 450 455 460
 Asn Asp Glu Ala Gly Phe Arg Lys Ala Met Asp Lys Ala Gly Lys Asn
 465 470 475 480
 Val Pro Leu Leu Ile Met Arg Arg Gly Asn Thr Leu Phe Ile Ala Leu
 485 490 495
 Asn Leu Gln

<210> 7

<211> 1110

<212> DNA

<213> Bacteria

<400> 7

atgttcaaaa aataccaata cctcgctttg gcagcactgt gtgccgcctc gctggcaggc	60
tgcgacaagg caggcagctt ttctgggtgcg gacaaaaaag aagcatcctt cgtagaacgc	120
atcgaaacaca ccaaagacga cggcagcgctc agtatgctgc tgcccgaactt tgcccgaactg	180
gttcaaagtg aaggtccggc agtcgtcaat attcaggcag ccccgcccc gcgcacccaa	240
aacggcagcg gcaatgccga aaacgattcc gaccgattg ccgacaacga cccgttctac	300
gaatttttca aacgcctcgt cccgaatatg cccgaaatcc cccaagaaga agcagatgac	360
ggcggattga acttcggttc gggcttcac atcagcaaag acggctacat cctgaccaat	420
acccacgtcg ttaccggcat gggcagtatc aaagtccctgc tcaacgacaa gcgcgaatat	480
accgccaac tcatcggttc ggatgtccaa tccgatgtcg cccttctgaa aatcgacgca	540
acggaagagc tgcccgtcgt caaaatcggc aatcccaaag atttgaaacc gggcgaatgg	600
gtcgccgcca tcggcgcgcc ctctgggttc gacaacagcg tgaccgcccg catcgtgtcc	660
gccaaaggca gaagcctgcc caacgaaagc tacacaccct tcatccaaac cgacgttgcc	720
atcaatccgg gcaactccgg cggcccgtcg ttcaacttaa aaggacaggt cgtcggcatc	780
aactcgcaaa tatacagccg cagcggcgga ttcatgggca ttctcttcgc catcccgatt	840
gacgttgcca tgaatgtcgc cgaacagctg aaaaacaccg gcaaagtcca acgcgacaa	900
ctgggctgta ttattcaaga agtatcctac ggtttggcac aatcgttcgg tttggacaaa	960
gccggcggcg cactgattgc caaaatcctg cccggcagcc ccgcagaacg tgccggcctg	1020
caggcgggcg acatcgtcct cagcctcgac ggcggagaaa tacgttcttc cggcgacctt	1080
cccggttatgg tcggcgccat tacgcccggga	1110

Asp Pro Phe Tyr Glu Phe Phe Lys Arg Leu Val Pro Asn Met Pro Glu
 100 105 110
 Ile Pro Gln Glu Glu Ala Asp Asp Gly Gly Leu Asn Phe Gly Ser Gly
 115 120 125
 Phe Ile Ile Ser Lys Asp Gly Tyr Ile Leu Thr Asn Thr His Val Val
 130 135 140
 Thr Gly Met Gly Ser Ile Lys Val Leu Leu Asn Asp Lys Arg Glu Tyr
 145 150 155 160
 Thr Ala Lys Leu Ile Gly Ser Asp Val Gln Ser Asp Val Ala Leu Leu
 165 170 175
 Lys Ile Asp Ala Thr Glu Glu Leu Pro Val Val Lys Ile Gly Asn Pro
 180 185 190
 Lys Asp Leu Lys Pro Gly Glu Trp Val Ala Ala Ile Gly Ala Pro Phe
 195 200 205
 Gly Phe Asp Asn Ser Val Thr Ala Gly Ile Val Ser Ala Lys Gly Arg
 210 215 220
 Ser Leu Pro Asn Glu Ser Tyr Thr Pro Phe Ile Gln Thr Asp Val Ala
 225 230 235 240
 Ile Asn Pro Gly Asn Ser Gly Gly Pro Leu Phe Asn Leu Lys Gly Gln
 245 250 255
 Val Val Gly Ile Asn Ser Gln Ile Tyr Ser Arg Ser Gly Gly Phe Met
 260 265 270
 Gly Ile Ser Phe Ala Ile Pro Ile Asp Val Ala Met Asn Val Ala Glu
 275 280 285
 Gln Leu Lys Asn Thr Gly Lys Val Gln Arg Gly Gln Leu Gly Val Ile
 290 295 300
 Ile Gln Glu Val Ser Tyr Gly Leu Ala Gln Ser Phe Gly Leu Asp Lys
 305 310 315 320
 Ala Gly Gly Ala Leu Ile Ala Lys Ile Leu Pro Gly Ser Pro Ala Glu
 325 330 335
 Arg Ala Gly Leu Gln Ala Gly Asp Ile Val Leu Ser Leu Asp Gly Gly
 340 345 350
 Glu Ile Arg Ser Ser Gly Asp Leu Pro Val Met Val Gly Ala Ile Thr
 355 360 365
 Pro Gly Lys Glu Val Ser Leu Gly Val Trp Arg Lys Gly Glu Glu Ile
 370 375 380
 Thr Ile Lys Val Lys Leu Gly Asn Ala Ala Glu His Ile Gly Ala Ser
 385 390 395 400
 Ser Lys Thr Asp Glu Ala Pro Tyr Thr Glu Gln Gln Ser Gly Thr Phe

```

gaatttttca aacgcctcgt cccgaatatg cccgaaatcc cccaagaaga agcagatgac      360
ggcggattga acttcggttc gggcttcacg atcagcaaag acggctacat cctgaccaat      420
acccacgtcg ttaccggcat gggcagtatc aaagtccctg tcaacgacaa gcgcgaatat      480
accgccaaac tcacgcgttc ggatgtccaa tccgatgtcg cccttctgaa aatcgacgca      540
acggaagagc tgcccgtcgt caaaatcggc aatcccaaag atttgaaacc gggcgaatgg      600
gtcgccgcca tcggcgcgcc cttcgggttc gacaacagcg tgaccgcggg catcgtgtcc      660
gccaaaggca gaagcctgcc caacgaaagc tacacaccct tcatccaaac cgacgttgcc      720
atcaatccgg gcaactccgg cggcccgtcg ttcaacttaa aaggacaggt cgtcggcatc      780
aactcgcaaa tatacagccg cagcggcgga ttcattgggca tttccttcgc catcccgatt      840
gacgttgcca tgaatgtcgc cgaacagctg aaaaacaccg gcaaagtcca acgcggacaa      900
ctgggcgtga ttattcaaga agtatcctac ggtttggtcac aatcgttcgg cttggacaaa      960
gccggcgggc cactgattgc caaaatcctg cccggcagcc ccgcagaacg tgccggcctg     1020
caggcggggc acatcgtcct cagcctcgac ggcgagaaaa tacgttcttc cggcgacctt     1080
cccgttatgg tcggcgccat tacgcgggga aaagaagtca gcctcggcgt atggcgcaaa     1140
ggcgaagaaa tcacaatcaa agtcaagctg ggcaacgccg ccgagcatat cggcgcatca     1200
tccaaaacag atgaagcccc ctacaccgaa cagcaatccg gtacgttctc ggtcgaatcc     1260
gcaggcatta cccttcagac acataccgac agcagcggcg gacacctcgt cgtcgtacgg     1320
gtttccgacg cggcagaacg cgcaggcttg aggcgcggcg acgaaattct tgccgtcggg     1380
caagtccccg tcaatgacga agccggtttc cgcaaagcta tggacaaggc aggcaaaaaac     1440
gtccccctgc tgatcatgcg ccgtggcaac acgctgttca tcgcattaaa cctgcaataa     1500

```

<210> 6

<211> 499

<212> PRT

<213> Bacteria

<400> 6

```

Met Phe Lys Lys Tyr Gln Tyr Leu Ala Leu Ala Ala Leu Cys Ala Ala
 1              5              10              15
Ser Leu Ala Gly Cys Asp Lys Ala Gly Ser Phe Phe Gly Ala Asp Lys
      20              25              30
Lys Glu Ala Ser Phe Val Glu Arg Ile Glu His Thr Lys Asp Asp Gly
      35              40              45
Ser Val Ser Met Leu Leu Pro Asp Phe Ala Gln Leu Val Gln Ser Glu
      50              55              60
Gly Pro Ala Val Val Asn Ile Gln Ala Ala Pro Ala Pro Arg Thr Gln
65              70              75              80
Asn Gly Ser Gly Asn Ala Glu Asn Asp Ser Asp Pro Ile Ala Asp Asn
      85              90              95

```

```

      290              295              300
Ile Gln Glu Val Ser Tyr Gly Leu Ala Gln Ser Phe Gly Leu Asp Lys
305              310              315              320
Ala Gly Gly Ala Leu Ile Ala Lys Ile Leu Pro Gly Ser Pro Ala Glu
      325              330              335
Arg Ala Gly Leu Gln Ala Gly Asp Ile Val Leu Ser Leu Asp Gly Gly
      340              345              350
Glu Ile Arg Ser Ser Gly Asp Leu Pro Val Met Val Gly Ala Ile Thr
      355              360              365
Pro Gly Lys Glu Val Ser Leu Gly Val Trp Arg Lys Gly Lys Glu Ile
      370              375              380
Thr Val Ala Val Lys Leu Gly Asn Ala Ser Glu Gln Thr Gly Ser Ser
385              390              395              400
Ser Glu Pro Asp Lys Ala Pro Tyr Ala Glu His Gln Ser Gly Thr Phe
      405              410              415
Ser Val Glu Ser Ala Gly Ile Thr Leu Gln Thr His Thr Asp Ser Ser
      420              425              430
Gly Gly Arg Leu Val Val Val Arg Val Ser Gly Ala Ala Glu Arg Ala
      435              440              445
Gly Leu Arg Arg Gly Asp Glu Ile Leu Ala Val Gly Gln Val Pro Val
      450              455              460
Asn Asp Glu Asp Gly Phe Arg Lys Ala Met Asp Lys Ala Gly Lys Asn
465              470              475              480
Val Pro Leu Leu Val Met Arg Arg Gly Asn Thr Leu Phe Ile Ala Leu
      485              490              495
Asn Leu Gln

```

<210> 5

<211> 1500

<212> DNA

<213> Bacteria

<400> 5

```

gtgtttcaaaa aataccaata cctcgctttg gcagcactgt gtgccgcctc gctggcaggc      60
tgcgacaagg caggcagctt ttctcgtgcg gacaaaaaag aagcatcctt cgtagaacgc      120
atcgaacaca ccaaagacga cggcagcgtc agtatgctgc tgcccgactt tgcccaactg      180
gttcaaagtg aaggtccggc agtcgtcaat attcaggcag ccccgcccc gcgcacccaa      240
aacggcagcg gcaatgccga aaacgattcc gacccgattg ccgacaacga cccgttctac      300

```

<400> 4

Met Phe Lys Lys Tyr Gln Tyr Leu Ala Leu Ala Ala Leu Cys Ala Ala
 1 5 10 15
 Ser Leu Ala Gly Cys Asp Lys Ala Gly Ser Phe Phe Gly Ala Asp Lys
 20 25 30
 Lys Glu Ala Ser Phe Val Glu Arg Ile Glu His Thr Lys Asp Asp Gly
 35 40 45
 Ser Val Ser Met Leu Leu Pro Asp Phe Val Gln Leu Val Gln Ser Glu
 50 55 60
 Gly Pro Ala Val Val Asn Ile Gln Ala Ala Pro Ala Pro Arg Thr Gln
 65 70 75 80
 Asn Gly Ser Gly Asn Ala Glu Thr Asp Ser Asp Pro Leu Ala Asp Ser
 85 90 95
 Asp Pro Phe Tyr Glu Phe Phe Lys Arg Leu Val Pro Asn Met Pro Glu
 100 105 110
 Ile Pro Gln Glu Glu Ala Asp Asp Gly Gly Leu Asn Phe Gly Ser Gly
 115 120 125
 Phe Ile Ile Ser Lys Asp Gly Tyr Ile Leu Thr Asn Thr His Val Val
 130 135 140
 Thr Gly Met Gly Ser Ile Lys Val Leu Leu Asn Asp Lys Arg Glu Tyr
 145 150 155 160
 Thr Ala Lys Leu Ile Gly Ser Asp Val Gln Ser Asp Val Ala Leu Leu
 165 170 175
 Lys Ile Asp Ala Thr Glu Glu Leu Pro Val Val Lys Ile Gly Asn Pro
 180 185 190
 Lys Asp Leu Lys Pro Gly Glu Trp Val Ala Ala Ile Gly Ala Pro Phe
 195 200 205
 Gly Phe Asp Asn Ser Val Thr Ala Gly Ile Val Ser Ala Lys Gly Arg
 210 215 220
 Ser Leu Pro Asn Glu Ser Tyr Thr Pro Phe Ile Gln Thr Asp Val Ala
 225 230 235 240
 Ile Asn Pro Gly Asn Ser Gly Gly Pro Leu Phe Asn Leu Lys Gly Gln
 245 250 255
 Val Val Gly Ile Asn Ser Gln Ile Tyr Ser Arg Ser Gly Gly Phe Met
 260 265 270
 Gly Ile Ser Phe Ala Ile Pro Ile Asp Val Ala Met Asn Val Ala Glu
 275 280 285
 Gln Leu Lys Asn Thr Gly Lys Val Gln Arg Gly Gln Leu Gly Val Ile

Asn Leu Gln

<210> 3

<211> 1500

<212> DNA

<213> Bacteria

<400> 3

```

gtgttcaaaa aataccaata cctcgctttg gcagcactgt gtgccgcctc gctggcaggc      60
tgcgacaaag cgggcagctt ttctgggtgcg gacaaaaaag aagcatcctt cgtagaacgc      120
atcgaaacaca ccaaagacga cggcagcgtc agtatgctgc tgcccgactt tgtccaactg      180
gttcaaagcg aaggcccggc agtcgtcaat attcaggcag ccccgcccc ggcgacccaa      240
aacggcagcg gcaatgccga aaccgattcc gaccgcttg ccgacagcga cccgttctac      300
gaatttttca aacgcctcgt ccggaatatg cccgaaatcc cccaagaaga agcagatgac      360
ggcggattga acttcggttc gggcttcac atcagcaaag acggctatat tctgaccaat      420
acgcacgtcg ttaccggcat gggcagtatc aaagtcctgc tcaacgacaa gcgcgaatat      480
accgccaaac tcacggttc ggatgtccaa tccgatgtcg cccttctgaa aatcgacgca      540
acggaagagc tgcccgctcg caaaatcggc aatcccaaag atttgaaacc gggcgaatgg      600
gtcgcgcgcca tcggcgcgcc ctctcggttc gacaacagcg tgaccgccgg catcgtgtcc      660
gccaaaggca gaagcctgcc caacgaaagc tacacaccct tcatccaaac cgacgttgcc      720
atcaatccgg gcaactccgg cggcccgtcg ttcaacctga aaggacaggt cgtcggcatc      780
aactcgcaaa tatacagccg cagcggcgga ttcatgggca ttcccttcgc catcccgatt      840
gacgttgcca tgaatgtcgc cgaacagctg aaaaacaccg gcaaagtcca acgcgacaaa      900
ctgggcgtga ttattcaaga agtatcctac ggtttggcac aatcgttcgg tttggacaaa      960
gccggcgggc cactgattgc caaaatcctg cccggcagcc ccgcagaacg tgccggcctg     1020
caggcggggc acatcgtcct cagcctcgac ggcgagaaaa tacgttcttc cggcgacctt     1080
cccgttatgg tcggcgccat tacgcccggg aaagaagtca gcctcggcgt atggcgcaaa     1140
ggtaaggaaa tcaccgttgc cgtcaaactg ggcaatgctt ccgaacaaac cggttcctcg     1200
tccgagccgg acaaagcccc ttatgccgaa caccaatccg gtacgttctc ggtcgaatcc     1260
gcaggcatta cccttcagac acataccgac agcagcggcg gacggcttgt cgtcgtgcgg     1320
gtttcggggg cggcagaacg cgcaggcttg aggcgcggcg acgaaatcct tgccgtcggg     1380
caagtccccg tcaatgacga agacggtttc cgcaaagcta tggacaaggc aggcaaaaac     1440
gtccccctgc tggatcatgc ccgtggcaac acgctgttca tcgcattaaa cctgcaataa     1500

```

<210> 4

<211> 499

<212> PRT

<213> Bacteria

180	185	190
Lys Asp Leu Lys Pro Gly Glu Trp Val Ala Ala Ile Gly Arg Pro Phe		
195	200	205
Gly Phe Asp Asn Ser Val Thr Ala Gly Ile Val Ser Ala Lys Gly Arg		
210	215	220
Ser Leu Pro Asn Glu Ser Tyr Thr Pro Phe Ile Gln Thr Asp Val Ala		
225	230	235
Ile Asn Pro Gly Asn Ser Gly Gly Pro Leu Phe Asn Leu Lys Gly Gln		240
245	250	255
Val Val Gly Ile Asn Ser Gln Ile Tyr Ser Arg Ser Gly Gly Phe Met		
260	265	270
Gly Ile Ser Phe Ala Ile Pro Ile Asp Val Ala Met Asn Val Ala Glu		
275	280	285
Gln Leu Lys Asn Thr Gly Lys Val Gln Arg Gly Gln Leu Gly Val Ile		
290	295	300
Ile Gln Glu Val Ser Tyr Gly Leu Ala Gln Ser Phe Gly Leu Asp Lys		
305	310	315
Ala Gly Gly Ala Leu Ile Ala Lys Ile Leu Pro Gly Ser Pro Ala Glu		320
325	330	335
Arg Ala Gly Leu Gln Ala Gly Asp Ile Val Leu Ser Leu Asp Gly Gly		
340	345	350
Glu Ile Arg Ser Ser Gly Asp Leu Pro Val Met Val Gly Ala Ile Thr		
355	360	365
Pro Gly Lys Glu Val Ser Leu Gly Val Trp Arg Lys Gly Lys Glu Ile		
370	375	380
Thr Val Ala Val Lys Leu Gly Asn Ala Ser Glu Gln Thr Gly Ser Ser		
385	390	395
Ser Glu Pro Asp Lys Ala Pro Tyr Ala Glu His Gln Ser Gly Thr Phe		400
405	410	415
Ser Val Glu Ser Ala Gly Ile Thr Leu Gln Thr His Thr Asp Ser Ser		
420	425	430
Gly Gly Arg Leu Val Val Val Arg Val Ser Gly Ala Ala Glu Arg Ala		
435	440	445
Gly Leu Arg Arg Gly Asp Glu Ile Leu Ala Val Gly Gln Val Pro Val		
450	455	460
Asn Asp Glu Asp Gly Phe Arg Lys Ala Met Asp Lys Ala Gly Lys Asn		
465	470	475
Val Pro Leu Leu Val Met Arg Arg Gly Asn Thr Leu Phe Ile Ala Leu		480
485	490	495

```

gccggcggcg cactgattgc caaaatcctg cccggcagcc ccgcagaacg tgccggcctg   1020
caggcggggcg acatcgctct cagcctcgac ggcggagaaa tacgtttctt cggcgacctt   1080
cccgttatgg tcggcgccat tacgcggga aaagaagtca gcctcggcgt atggcgcaaa   1140
ggtaaggaaa tcaccgttgc cgtcaaaactg ggcaatgctt ccgaacaaac cggttcctcg   1200
tccgagccgg acaaagcccc ttatgccgaa caccaatccg gtacgtttctc ggtcgaatcc   1260
gcaggcatta cccttcagac acataccgac agcagcggcg gacggcttgt cgtcgtgcgg   1320
gtttcggggg cggcagaacg cgcaggcttg aggcgcggcg acgaaatcct tgccgtcggg   1380
caagtccccg tcaatgacga agacggtttc cgcaaagcta tggacaaggc aggcaaaaac   1440
gtccccctgc tggtcatgcg ccgtggcaac acgctgttca tcgcattaaa cctgcaataa   1500

```

<210> 2

<211> 499

<212> PRT

<213> Bacteria

<400> 2

```

Met Phe Lys Lys Tyr Gln Tyr Leu Ala Leu Ala Ala Leu Cys Ala Ala
 1              5              10              15
Ser Leu Ala Gly Cys Asp Lys Ala Gly Ser Phe Phe Gly Ala Asp Lys
      20              25              30
Lys Glu Ala Ser Phe Val Glu Arg Ile Glu His Thr Lys Asp Asp Gly
      35              40              45
Ser Val Ser Met Leu Leu Pro Asp Phe Val Gln Leu Val Gln Ser Glu
      50              55              60
Gly Pro Ala Val Val Asn Ile Gln Ala Ala Pro Ala Pro Arg Thr Gln
      65              70              75              80
Asn Gly Ser Gly Asn Ala Glu Thr Asp Ser Asp Pro Leu Ala Asp Ser
      85              90              95
Asp Pro Phe Tyr Glu Phe Phe Lys Arg Leu Val Pro Asn Met Pro Glu
      100             105             110
Ile Pro Gln Glu Glu Ala Asp Asp Gly Gly Leu Asn Phe Gly Ser Gly
      115             120             125
Phe Ile Ile Ser Lys Asp Gly Tyr Ile Leu Thr Asn Thr His Val Val
      130             135             140
Thr Gly Met Gly Ser Ile Lys Val Leu Leu Asn Asp Lys Arg Glu Tyr
      145             150             155             160
Thr Ala Lys Leu Ile Gly Ser Asp Val Gln Ser Asp Val Ala Leu Leu
      165             170             175
Lys Ile Asp Ala Thr Glu Glu Leu Pro Val Val Lys Ile Gly Asn Pro

```

SEQUENCE LISTING

<110> SmithKline Beecham Biologicals S.A.

<120> Novel compounds

<130> BM45307

<160> 16

<170> FastSEQ for Windows Version 3.0

<210> 1

<211> 1500

<212> DNA

<213> Bacteria

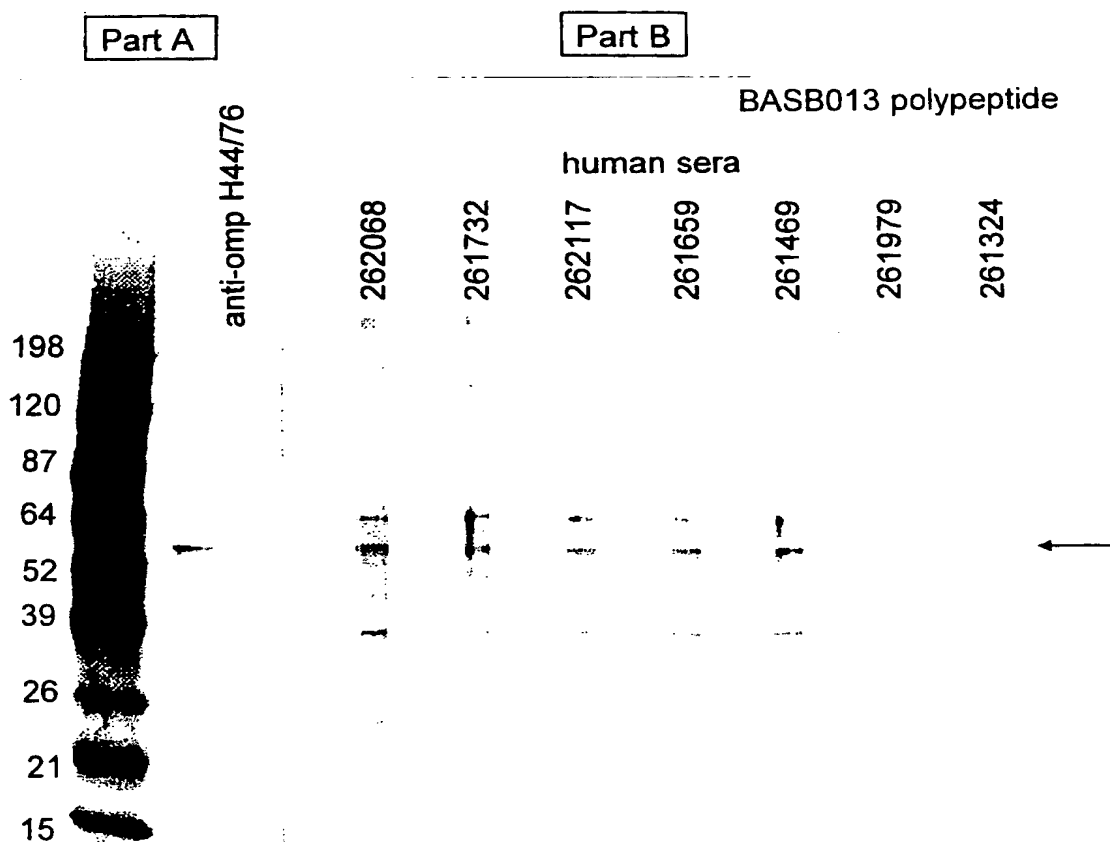
<400> 1

gtgttcaaaa aataccaata cctcgctttg gcagcactgt gtgccgcctc gctggcaggc	60
tgcgacaaaag ccggcagctt ttctgggtgcg gacaaaaaag aagcctcctt cgtagaacgc	120
atcgaacaca ccaaagacga cggcagcgtc agtatgctgc tgcccgaactt tgtccaactg	180
gttcaaagcg aaggcccggc agtcgtcaat attcaggcag cccccgcccc gcgcacccaa	240
aacggcagcg gcaatgccga aaccgattcc gaccgcttg ccgacagcga ccggttctac	300
gaatttttca aacgcctcgt cccgaatatg cccgaaatcc cccaagaaga agcagatgac	360
ggcggattga acttcgggttc gggcttcac atcagcaaag acggctatat tctgaccaat	420
acgcacgtcg ttaccggcat gggcagtatc aaagtcctgc tcaacgacaa gcgcgaatat	480
accgccaaac tcatcggttc ggatgtccaa tccgatgtcg cccttctgaa aatcgacgca	540
acggaagagc tgcccgtcgt caaaatcggc aatcccaaag atttgaaacc gggcgaatgg	600
gtcgccgcca tcggggcgcc cttcggcttc gacaacagcg tgaccgccgg catcgtgtcc	660
gccaaaggca gaagcctgcc caacgaaagc tacacaccct tcatccaaac cgacgttgcc	720
atcaatccgg gcaactccgg cggcccgtcg ttcaacctga aaggacaggt cgtcggcatc	780
aactcgcaaa tatacagccg cagcggcgga ttcatgggca ttctcttcgc catcccgatt	840
gacgttgcca tgaatgtcgc cgaacagctg aaaaacaccg gcaaagtcca acgcggacaa	900
ctgggcgtga ttattcaaga agtatcctac ggtttggcac aatcgttcgg ttggacaaa	960

9/9

Fig.4

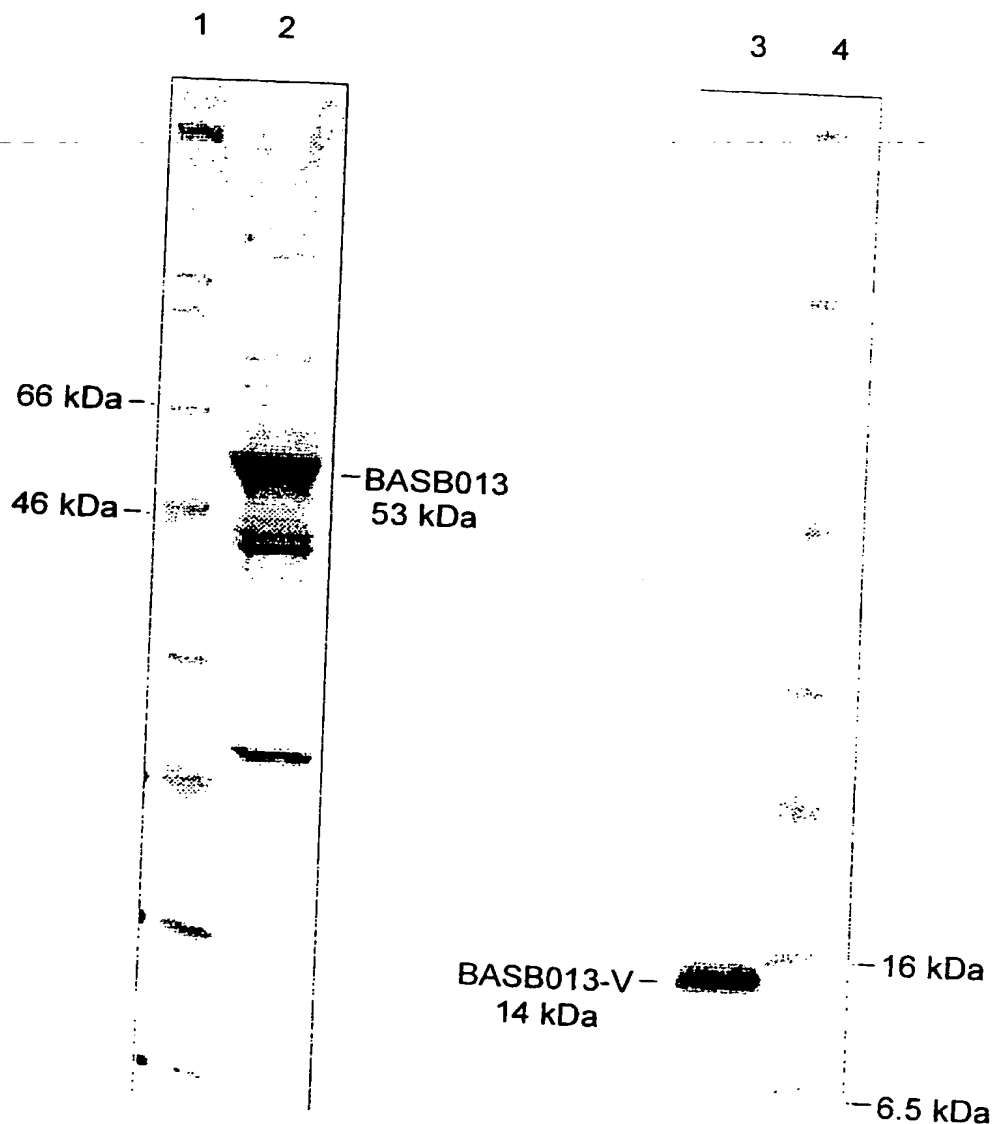
Anti-BASB013 antibodies in human convalescent sera (part B) and mice immunized with Outer Membrane Proteins of H44/76 *Neisseria meningitidis* cells (part A).



8/9

Fig.3

Expression and purification of recombinant BASB013 forms. Substantially purified proteins were separated on a 4-20% gradient polyacrylamide gel (NOVEX) under SDS-PAGE conditions and stained with Coomassie Blue R250. The sample loaded on the gel corresponded to molecular weight marker (lanes 1 and 4) and protein fractions enriched (more than 80%) in BASB013 (lane2) and BASB013-V (lane3).



7/9

 * 320 * 340 *

Seqid2: LGVIIQEVSYG LAQSFGLDKAGGALIAKILPGSPAERAGLQAGDIVLSLD:350

Seqid4::350

Seqid6::350

 360 * 380 * 400

Seqid2: GGEIRSSGDLPMVGAITPGKEVSLGVWRKGKEITVAVKLG NASEQTGSS:400

Seqid4::400

Seqid6:E...IK.....A..HI...A.:400

 * 420 * 440 *

Seqid2: SEPDKAPYAEHQSGTFSVESAGITLQTHTDSSGGRLVVVRVSGAAERAGL:450

Seqid4::450

Seqid6: ..KT..E.....T..Q.....H.....D.....:450

 460 * 480 *

Seqid2: RRGDEILAVGQVPVNDEDGFRKAMDKAGKNVPLLVMRRGNTLFIALNLQ:499

Seqid4::499

Seqid6:A.....I.....:499

6/9

Fig.2

Alignment of the BASB013 polypeptide sequences. Identity to SeqID No:2 is indicated by a dot.

	*	20	*	40	*
Seqid2:	MFKKYQYLALAALCAASLAGCDKAGSFFGADKKEASFVERIEHTKDDGSV:	50			
Seqid4:	50			
Seqid6:	50			

	60	*	80	*	100
Seqid2:	SMLLPDFVQLVQSEGPVVNIQAAPAPRTQNGSGNAETDSDPLADSDPFY:	100			
Seqid4:	100			
Seqid6:	A.....N.....I.....N.....	100			

	*	120	*	140	*
Seqid2:	EFFKRLVPNMPEIPQEEADDGGLNFGSGFIISKDGYILTNTHVVTGMGSI:	150			
Seqid4:	150			
Seqid6:	150			

	160	*	180	*	200
Seqid2:	KVLLNDKREYTAKLIGSDVQSDVALLKIDATEELPVVKIGNPKDLKPGEW:	200			
Seqid4:	200			
Seqid6:	200			

	*	220	*	240	*
Seqid2:	VAAIGRPFGFDNSVTAGIVSAKGRSLPNESYTPFIQTDVAINPGNSGGPL:	250			
Seqid4:	A.....	250			
Seqid6:	A.....	250			

	260	*	280	*	300
Seqid2:	FNLKGQVVGINSQIYSRSGGFMGISFAIPIDVAMNVAEQLKNTGKVQRGQ:	300			
Seqid4:	300			
Seqid6:	300			

5/9

 * 1320 * 1340 *

Seqid1:GACGGCTTGTCTGTCGTGCGGGTTTCGGGGGCGGCAGAACGCGCAGGCTTG:1350

Seqid3:.....:1350

Seqid5:....AC.....C.....A.....C..AC.....:1350

 1360 * 1380 * 1400

Seqid1:AGGCGCGGCGACGAAATCCTTGCCGTGGGCAAGTCCCCGTCAATGACGA:1400

Seqid3:.....:1400

Seqid5:.....T.....:1400

 * 1420 * 1440 *

Seqid1:AGACGGTTTCCGCAAAGCTATGGACAAGGCAGGCAAAAACGTCCCCCTGC:1450

Seqid3:.....:1450

Seqid5:....C.....:1450

 1460 * 1480 * 1500

Seqid1:TGGTCATGCGCCGTGGCAACACGCTGTTTCATCGCATTAAACCTGCAATAA:1500

Seqid3:.....:1500

Seqid5:....A.....:1500

4/9

* 1020 * 1040 *

Seqid1: CCGCAGAACGTGCCGGCCTGCAGGCGGGCGACATCGTCCTCAGCCTCGAC:1050
 Seqid3::1050
 Seqid5::1050

1060 * 1080 * 1100

Seqid1: GCGGAGAAATACGTTCTTCCGGCGACCTTCCCGTTATGGTCGGCGCCAT:1100
 Seqid3::1100
 Seqid5::1100

* 1120 * 1140 *

Seqid1: TACGCCGGGAAAAGAAGTCAGCCTCGGCGTATGGCGCAAAGGTAAGGAAA:1150
 Seqid3::1150
 Seqid5:CG.A.....:1150

1160 * 1180 * 1200

Seqid1: TCACCGTTGCCGTCAAAGTGGGCAATGCTTCCGAACAAACCGGTTCTCG:1200
 Seqid3::1200
 Seqid5:AA...CAAA.....G.....C...CG.....G...T..T....CG..A...A:1200

* 1220 * 1240 *

Seqid1: TCCGAGCCGGACAAAGCCCCCTTATGCCGAACACCAATCCGGTACGTTCTC:1250
 Seqid3::1250
 Seqid5:A..AA..A...TG.....C...CA.....G.....:1250

1260 * 1280 * 1300

Seqid1: GGTGGAATCCGCAGGCATTACCTTCAGACACATACCGACAGCAGCGGCG:1300
 Seqid3::1300
 Seqid5::1300

SUBSTITUTE SHEET (RULE 26)

3/9

660 * 680 * 700
 Seqid1:CATCGTGTCGCGCCAAAGGCAGAAGCCTGCCCAACGAAAGCTACACACCCT:700
 Seqid3:.....:700
 Seqid5:.....:700

* 720 * 740 *
 Seqid1:TCATCCAAACCGACGTTGCCATCAATCCGGGCAACTCCGGCGGCCCGCTG:750
 Seqid3:.....:750
 Seqid5:.....:750

760 * 780 * 800
 Seqid1:TTCAACCTGAAAGGACAGGTCGTCGGCATCAACTCGCAAATATACAGCCG:800
 Seqid3:.....:800
 Seqid5:.....T.A.....:800

* 820 * 840 *
 Seqid1:CAGCGGCGGATTTCATGGGCATTTCTTCGCCATCCCGATTGACGTTGCCA:850
 Seqid3:.....:850
 Seqid5:.....:850

860 * 880 * 900
 Seqid1:TGAATGTCGCCGAACAGCTGAAAAACACCGGCAAAGTCCAACGCGGACAA:900
 Seqid3:.....:900
 Seqid5:.....:900

* 920 * 940 *
 Seqid1:CTGGGCGTGATTATTCAAGAAGTATCCTACGGTTTGGCACAATCGTTCGG:950
 Seqid3:.....:950
 Seqid5:.....:950

960 * 980 * 1000
 Seqid1:TTTGACAAAGCCGGCGGCGCACTGATTGCCAAAATCCTGCCCGGCAGCC:1000
 Seqid3:.....:1000
 Seqid5:.....:1000

2/9

* 320 * 340 *

Seqid1:GAATTTTCAAACGCCTCGTCCCGAATATGCCCGAAATCCCCCAAGAAGA:350
 Seqid3:.....:350
 Seqid5:.....:350

360 * 380 * 400

Seqid1:AGCAGATGACGGCGGATTGAACTTCGGTTCGGGCTTCATCATCAGCAAAG:400
 Seqid3:.....:400
 Seqid5:.....:400

* 420 * 440 *

Seqid1:ACGGCTATATTCTGACCAATACGCACGTCGTTACCGGCATGGGCAGTATC:450
 Seqid3:.....:450
 Seqid5:.....C...T.....C.....:450

460 * 480 * 500

Seqid1:AAAGTCCTGCTCAACGACAAGCGCGAATATACGCCAAACTCATCGGTTC:500
 Seqid3:.....:500
 Seqid5:.....:500

* 520 * 540 *

Seqid1:GGATGTCCAATCCGATGTCGCCCTTCTGAAAATCGACGCAACGGAAGAGC:550
 Seqid3:.....:550
 Seqid5:.....:550

560 * 580 * 600

Seqid1:TGCCCGTCGTCAAATCGGCAATCCCAAAGATTTGAAACCGGGCGAATGG:600
 Seqid3:.....:600
 Seqid5:.....:600

* 620 * 640 *

Seqid1:GTCGCCGCATCGGGCGGCCCTTCGGCTTCGACAACAGCGTGACCGCCGG:650
 Seqid3:.....CGC.....:650
 Seqid5:.....CGC.....:650

SUBSTITUTE SHEET (RULE 26)

1/9

Fig.1

Alignment of the BASB013 polynucleotide sequences. Identity to SeqID No:1 is indicated by a dot.

```

          *           20           *           40           *
Seqid1:GTGTTCAAAAAATACCAATACCTGCTTTGGCAGCACTGTGTGCCGCCTC:050
Seqid3:.....:050
Seqid5:.....:050

```

```

          60           *           80           *           100
Seqid1:GCTGGCAGGCTGCGACAAAGCCGGCAGCTTTTCGGTGGGACAAAAAAG:100
Seqid3:.....:100
Seqid5:.....G...A.....:100

```

```

          *           120           *           140           *
Seqid1:AAGCATCCTTCGTAGAACGCATCGAACACACCAAAGACGACGGCAGCGTC:150
Seqid3:.....:150
Seqid5:.....:150

```

```

          160           *           180           *           200
Seqid1:AGTATGCTGCTGCCCCGACTTTGTCCAAGTCAAGCGAAGGCCCGGC:200
Seqid3:.....:200
Seqid5:.....C.....T.....T.....:200

```

```

          *           220           *           240           *
Seqid1;AGTCGTCAATATTCAGGCAGCCCCCGCCCCGCGACCCAAAACGGCAGCG:250
Seqid:3.....:250
Seqid5:.....:250

```

```

          260           *           280           *           300
Seqid1:GCAATGCCGAAACCGATTCCGACCCGCTTGCCGACAGCGACCCGTTCTAC:300
Seqid3:.....:300
Seqid5:.....A.....A.....A.....:300

```

23. The vaccine composition according to either one of claims 21 or 22 wherein said composition comprises at least one other *Neisseria meningitidis* antigen.
24. An antibody immunospecific for the polypeptide or immunological fragment as
5 claimed in any one of claims 1 to 6.
25. A method of diagnosing a *Neisseria meningitidis* infection, comprising identifying a polypeptide as claimed in any one of claims 1 - 6, or an antibody that is immunospecific for said polypeptide, present within a biological sample from an animal suspected of
10 having such an infection.
26. Use of a composition comprising an immunologically effective amount of a polypeptide as claimed in any one of claims 1 - 6 in the preparation of a medicament for use in generating an immune response in an animal.
15
27. Use of a composition comprising an immunologically effective amount of a polynucleotide as claimed in any one of claims 7 - 16 in the preparation of a medicament for use in generating an immune response in an animal.
- 20 28. A therapeutic composition useful in treating humans with *Neisseria meningitidis* disease comprising at least one antibody directed against the polypeptide of claims 1 - 6 and a suitable pharmaceutical carrier.

hybridization conditions with a labeled probe having the sequence of SEQ ID NO:1 or a fragment thereof.

5 17. An expression vector or a recombinant live microorganism comprising an isolated polynucleotide according to any one of claims 7 - 16.

18. A host cell comprising the expression vector of claim 17 or a subcellular fraction or a membrane of said host cell expressing an isolated polypeptide comprising an amino acid sequence that has at least 85% identity to the amino acid sequence selected from the group
10 consisting of: SEQ ID NO:4, SEQ ID NO:6.

19. A process for producing a polypeptide comprising an amino acid sequence that has at least 85% identity to the amino acid sequence selected from the group consisting of: SEQ ID NO:4, SEQ ID NO:6 comprising culturing a host cell of claim 18 under conditions
15 sufficient for the production of said polypeptide and recovering the polypeptide from the culture medium.

20. A process for expressing a polynucleotide of any one of claims 7 – 16 comprising transforming a host cell with the expression vector comprising at least one of said
20 polynucleotides and culturing said host cell under conditions sufficient for expression of any one of said polynucleotides.

21. A vaccine composition comprising an effective amount of the polypeptide of any one of claims 1 to 6 and a pharmaceutically acceptable carrier.

25

22. A vaccine composition comprising an effective amount of the polynucleotide of any one of claims 7 to 16 and a pharmaceutically effective carrier.

entire coding region; or a nucleotide sequence complementary to said isolated polynucleotide.

9. An isolated polynucleotide which comprises a nucleotide sequence which has at least
5 85% identity to that of SEQ ID NO: 3, 5 over the entire length of SEQ ID NO:3, 5
respectively; or a nucleotide sequence complementary to said isolated polynucleotide.

10. The isolated polynucleotide as claimed in any one of claims 7 to 9 in which the
identity is at least 95% to SEQ ID NO: 3, 5.

10

11. An isolated polynucleotide comprising a nucleotide sequence encoding the polypeptide
of SEQ ID NO:4, SEQ ID NO:6.

12. An isolated polynucleotide comprising the polynucleotide of SEQ ID NO:3, SEQ ID
15 NO:5.

13. An isolated polynucleotide comprising a nucleotide sequence encoding the polypeptide
of SEQ ID NO:4, SEQ ID NO:6 obtainable by screening an appropriate library under
stringent hybridization conditions with a labeled probe having the sequence of SEQ ID
20 NO:3, SEQ ID NO:5 or a fragment thereof.

14. An isolated polynucleotide comprising a nucleotide sequence encoding the polypeptide
of SEQ ID NO:2.

25 15. An isolated polynucleotide comprising the polynucleotide of SEQ ID NO:1.

16. An isolated polynucleotide comprising a nucleotide sequence encoding the polypeptide
of SEQ ID NO:2 obtainable by screening an appropriate library under stringent

CLAIMS:

1. An isolated polypeptide comprising an amino acid sequence which has at least 85%
5 identity to the amino acid sequence selected from the group consisting of: SEQ ID NO:4, SEQ ID NO:6.
2. An isolated polypeptide as claimed in claim 1 in which the amino acid sequence has at least 95% identity to the amino acid sequence selected from the group consisting of: SEQ
10 ID NO:4, SEQ ID NO:6.
3. The polypeptide as claimed in claim 1 comprising the amino acid sequence selected from the group consisting of: SEQ ID NO:4, SEQ ID NO:6.
- 15 4. An isolated polypeptide of SEQ ID NO:4, SEQ ID NO:6.
5. An isolated polypeptide of SEQ ID NO:2.
6. An immunogenic fragment of the polypeptide as claimed in any one of claims 1 to 5 in
20 which the immunogenic activity of said immunogenic fragment is substantially the same as the polypeptide of SEQ ID NO:4, SEQ ID NO:6.
7. An isolated polynucleotide comprising a nucleotide sequence encoding a polypeptide that has at least 85% identity to the amino acid sequence of SEQ ID NO: 4, 6 over the
25 entire length of SEQ ID NO: 4, 6 respectively; or a nucleotide sequence complementary to said isolated polynucleotide.
8. An isolated polynucleotide comprising a nucleotide sequence that has at least 85% identity to a nucleotide sequence encoding a polypeptide of SEQ ID NO: 4, 6 over the

INDICATIONS RELATING TO DEPOSITED MICROORGANISM
OR OTHER BIOLOGICAL MATERIAL

(PCT Rule 13bis)

A. The indications made below relate to the deposited microorganism or other biological material referred to in the description on page <u>56</u> , line <u>2-21</u>	
B. IDENTIFICATION OF DEPOSIT Further deposits are identified on an additional sheet <input type="checkbox"/>	
Name of depositary institution AMERICAN TYPE CULTURE COLLECTION	
Address of depositary institution (including postal code and country) 10801 UNIVERSITY BLVD, MANASSAS, VIRGINIA 20110-2209, UNITED STATES OF AMERICA	
Date of deposit 22 June 1997 (22/06/97)	Accession Number 13090
C. ADDITIONAL INDICATIONS (leave blank if not applicable) This information is continued on an additional sheet <input type="checkbox"/>	
In respect of those designations where a European Patent is sought, a sample of the deposited microorganism will be made available until the publication of the mention of the grant of the European Patent or until the date on which the application has been refused or withdrawn, only by issue of such a sample to an expert nominated by the person requesting the sample.	
D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE (if the indications are not for all designated States)	
E. SEPARATE FURNISHING OF INDICATIONS (leave blank if not applicable)	
The indications listed below will be submitted to the International Bureau later (specify the general nature of the indications e.g., "Accession Number of Deposit")	
<div style="border: 1px solid black; padding: 5px;"><div style="text-align: center; font-weight: bold; font-size: small;">For receiving Office use only</div><div style="display: flex; align-items: center; margin-bottom: 10px;"><input type="checkbox"/> This sheet was received with the international application</div><div style="border-top: 1px solid black; height: 40px; margin-top: 10px;"></div></div>	<div style="border: 1px solid black; padding: 5px;"><div style="text-align: center; font-weight: bold; font-size: small;">For International Bureau use only</div><div style="display: flex; align-items: center; margin-bottom: 10px;"><input type="checkbox"/> This sheet was received by the International Bureau on:</div><div style="border-top: 1px solid black; height: 40px; margin-top: 10px;"></div></div>